

Development of Linear Scaling *ab initio*
Methods based on Electron Density Matrices

Entwicklung linear skalierender *ab initio*
Methoden basierend auf
Elektronendichtematrizen

DISSERTATION

der Fakultät für Chemie und Pharmazie
der Eberhard-Karls-Universität Tübingen

zur Erlangung des Grades eines Doktors
der Naturwissenschaften

2006

vorgelegt von

Jörg Kußmann

Tag der mündlichen Prüfung: 14.06.2006

Dekan: Prof. Dr. S. Laufer

1. Berichterstatter: Prof. Dr. C. Ochsenfeld

2. Berichterstatter: Prof. Dr. H. Oberhammer

3. Berichterstatter: Prof. Dr. W. Thiel

Herrn Prof. Dr. C. Ochsenfeld danke ich für die Bereitstellung des interessanten Themas, seine Hilfsbereitschaft und die Möglichkeit vieler Diskussionen. Weiterhin möchte ich mich für sein Vertrauen und den Freiraum zur Entfaltung eigener Ideen bedanken.

Herrn Prof. Dr. H. Oberhammer danke ich für die Übernahme des Koreferats sowie Herrn Prof. Dr. W. Thiel für das dritte Gutachtung zu meiner Arbeit.

Für die gute Arbeitsatmosphäre sowie für ihre Hilfs- und Diskussionsbereitschaft danke ich allen gegenwärtigen und ehemaligen Kollegen: Bernd Doser, Marc Fuhrmans, Dr. Dietrich Hoffmann, Dr. Felix Koziol, Daniel Lambrecht, Martin Löffler, Dr. Alexander Neugebauer, Hella Riede, Benedikt Schmitt, Sabine Schweizer, Chris Sumowski, Mona Wambach und Jan Zienau. Weiterhin möchte ich Lothar Braun für seine Hilfe in Computer-Fragen danken.

Meinen Eltern, Rita und Rolf Kussmann, möchte ich für die stete Unterstützung und Hilfe jeglicher Art danken.

Oh, come with old Khayyám, and leave the wise
To talk; one thing is certain, that life flies;
One thing is certain, and the rest is lies;
The flower that once has blown for ever dies.

—*Rubáiyat of Omar Khayyám*

*Meinen Eltern
Rita und Rolf*

Contents

1	Introduction	5
2	Theory	9
2.1	Molecular (Non-Relativistic) Hamiltonian	9
2.2	Born-Oppenheimer Approximation	10
2.3	Hartree-Fock Theory	11
2.4	Linear Scaling Formation of the Fock Matrix	14
2.5	The Electron Density	15
2.5.1	1-particle Reduced Density Function of a Single Slater Determinant: Fock-Dirac Density	17
2.5.2	Properties of the Fock-Dirac Density	19
2.5.2.1	Purification Transformation	21
2.5.3	Structure of the Fock-Dirac Density	23
2.6	Electron Correlation	24
2.6.1	Pair Functions in HF-Theory	25
2.6.2	Expansion of a Many-Electron Wave Function in the Basis of One- Electron Functions: Full-CI Wave Functions	26
2.6.3	Density Functional Theory	28
2.6.3.1	Hohenberg-Kohn Theorems	29
2.6.3.2	Kohn-Sham-DFT	30
2.6.3.3	Time-Dependent Systems: Runge-Gross Theorem	31
2.6.3.4	Formation of Exchange-Correlation Matrices	32
3	Linear Scaling Density Matrix-based Methods for Static and Dynamic Properties	35

3.1	Perturbative Expansion of the Energy	36
3.2	First Order Derivatives	37
3.3	Properties of the First-Order Perturbed Density	38
3.4	Frequency-Dependent Perturbations and Properties of the First and Second Order Transition Densities	39
3.5	MO-based CPSCF/TDSCF Equations	42
3.5.1	Dynamic Properties: MO-TDSCF	43
3.5.1.1	First-Order MO-TDSCF Equations	44
3.5.1.2	Second-Order MO-TDSCF Equations	45
3.5.1.3	Wigner (2n+1) Rule	46
3.5.2	Static Second-Order Properties: First-Order CPSCF	47
3.6	Linear Scaling Methods for Static Second Order Properties: First Order Density Matrix-based CPSCF	47
3.6.1	Linear Response Equations to a Static Perturbation	48
3.6.1.1	A Brief Review of the Quadratically Convergent Density Matrix-based SCF Energy Minimization	48
3.6.1.2	D-CPSCF Scheme Derived from D-QCSCF	50
3.6.1.3	An Alternative Derivation of D-CPSCF	54
3.6.1.4	Comparison with Other CPSCF Algorithms	57
3.6.2	Nuclear Magnetic Shielding Tensor	58
3.6.2.1	Molecular Hamiltonian in the Presence of an External Mag- netic Field — The Gauge-Origin Problem	59
3.6.2.2	Explicit Expressions for the Different Terms in the Pertur- bative Expansion	61
3.6.2.3	GIAO Kohn-Sham DFT	63
3.6.3	Implementational Details	64
3.6.3.1	Linear Equation Solver	64
3.6.3.2	Integral Engines	65
3.6.4	Applications of D-GIAO-CPSCF	66
3.7	Linear Scaling Methods for Second and Third Order Properties: First and Second Order Density Matrix-based TDSCF	71
3.7.1	Linear Response Equations for a Dynamic Perturbation	71
3.7.1.1	Linear Response of the Exchange-Correlation Functional	72

3.7.1.2	Molecular Polarizability – Illustrative Examples	73
3.7.2	Quadratic Response Equations for a Dynamic Perturbation	76
3.7.2.1	Initial Guess for the Second Order Transition Density	77
3.7.2.2	Quadratic Response of the Exchange-Correlation Functional	78
3.7.2.3	Determination of Hyperpolarizabilities Exploiting Wigner’s $(2n + 1)$ Rule	79
3.7.2.4	First Molecular Hyperpolarizabilities – Illustrative Examples	83
4	Analysis of Sparse Algebra Routines within the D-GIAO-HF Algorithm	87
4.1	Screening in Sparse Matrix Multiplications	88
4.2	Behavior and Stability of Sparse Algebra Routines within an Iterative Process	90
4.3	The Influence of Truncating Matrices	92
4.4	Comparison with Standard Library Routines	94
5	Linear Scaling Quantum Monte Carlo Algorithms for the Local Energy with Rigorously Controllable Error Bounds	97
5.1	Stochastic Foundations	98
5.1.1	Markov Chain Monte Carlo: The Metropolis Algorithm	101
5.2	Variational Quantum Monte Carlo	103
5.2.1	Importance Sampling by Langevin-type Fictitious Dynamics	105
5.3	Diffusion Quantum Monte Carlo	106
5.4	Trial Wave Functions in Quantum Monte Carlo	109
5.4.1	The Slater Determinant	111
5.4.2	The Correlation Factor	112
5.5	N -Particle Density Matrix-based Variational Quantum Monte Carlo	115
5.5.1	Illustrative Calculations	118
5.6	N -Particle Density Matrix-based Diffusion Quantum Monte Carlo	122
5.6.1	Illustrative Calculations	125
6	Conclusions and Perspectives	129
	Bibliography	131
	List of Figures	141

List of Tables	145
A Derivatives of the Exchange-Correlation Potential	147
A.1 Local Spin Density Approximation	148
A.2 Generalized Gradient Approximation	148
B Figures	155
C Tables	177
D Atomic Units	183
E Abbreviations and Symbols	185
Publications	191
Academic Teachers	193
Curriculum vitae	195

Chapter 1

Introduction

In the last decades ab initio methods have become a standard tool in chemistry, biochemistry, and physics. Their reliability and accuracy in the prediction of properties of molecular and solid-state systems have been proven in many applications. Nevertheless, the size of treatable systems was limited to the hundred atoms region even for the less demanding Hartree-Fock (HF) and Kohn-Sham density functional (KS-DFT) theory because of their at least cubic scaling behavior $\mathcal{O}(M^3)$ with system size M , i.e. the computational effort is eight times larger if the size of the system is doubled. Ab initio methods of course profit by the fast evolution of computer technology, but their application to larger systems is primarily hampered by the unfavorable scaling behavior. Thus the development of linear scaling methods is a focus of quantum chemistry in order to expand the spectrum of treatable molecular systems.

With the development of $\mathcal{O}(M)$ algorithms for evaluating the two-electron contributions to the Fock or Kohn-Sham matrix (e.g. Refs. [1–10]) of systems with a significant HOMO-LUMO gap, the applicability has been extended to the thousand atoms region. For larger systems the cubic scaling behavior of the linear algebra routines for e.g. diagonalizing the Fock matrix starts to dominate the computational effort. In these cases one can employ density matrix-based diagonalization alternatives [11–15], where only local quantities enter and thus the application of $\mathcal{O}(M)$ sparse algebra routines becomes possible.

The aim of the present work is the development of new linear scaling methods for the calculation of molecular properties and quantum Monte Carlo (QMC) energies by reformulations in terms of electron density matrices. In the first part, an overall linear scaling method for predicting static second order properties is presented. We applied our new

coupled-perturbed self-consistent field (CPSCF) scheme to solve directly for the perturbed densities within the NMR chemical shieldings calculation, which are extremely important for attaining structural and dynamical insights into chemical systems. The calculation of NMR chemical shifts at the HF or KS-DFT level of theory is routine nowadays and has proven to yield reliable and accurate results in many cases [16–25]. However, the cubic scaling behavior of the standard MO-based schemes prevents the treatment of larger systems. In the present work a density matrix-based coupled perturbed self-consistent field (D-CPSCF) method is presented, which allows in combination with $\mathcal{O}(M)$ two-electron integral contractions to achieve an overall linear scaling behavior. In this way, the spectrum of treatable systems is extended from the hundred atoms region to molecules containing thousand and more atoms. Apart from exemplary calculations to prove the $\mathcal{O}(M)$ scaling behavior of the new method, first applications for studying solid- and solution-state systems are presented.

The second part of this thesis treats the calculation of frequency-dependent molecular polarizabilities and first hyperpolarizabilities [26, 27]. The importance of theoretical methods for the calculation of linear and non-linear optical properties grows with the interest in corresponding materials for optical devices. The efficiency of such devices like optical fibres, optical frequency converters, electro-optical modulators, thermo-optical switches etc. but also liquid crystals for TFT displays is given by the polarizabilities and hyperpolarizabilities as first objectives. In combination with ab initio calculations, which allow the detailed investigation of different influences on these effects, a carefully directed improvement of the optical properties can be possible which opens a new path to intelligent materials design. In addition to the traditional inorganic optical devices there is a great interest in organic materials nowadays, whose non-linear properties mainly originate from the change of the polarizability of π -electrons and so exhibit evidently faster optical effects. For such systems with a naturally local electronic structure the presented density matrix-based time-dependent self-consistent field (D-TDSCF) method enables an overall linear scaling behavior.

The final part of this work is focused on the development of a linear scaling method for the local energy in variational (VQMC) and diffusion (DQMC) quantum Monte Carlo [28, 29]. In recent years an increasing interest in quantum Monte Carlo methods was noticeable since they combine high accuracy and a favorable scaling behavior as compared to the more difficult "Post Hartree-Fock" methods like for example perturbation theory or

the coupled cluster (CC) approximation. In this work a QMC method with linear scaling effort in the computation of the local energy has been developed by a reformulation of the corresponding equations in the basis of the N -particle density matrix (N -PDM). Apart from the derivation and implementation of the new N -PDM VQMC and N -PDM FN-DQMC equations, first tests of their performance are shown for a series of linear alkanes.

Chapter 2

Theory

2.1 Molecular (Non-Relativistic) Hamiltonian

The central equation of non-relativistic quantum chemistry is the time-dependent Schrödinger equation

$$\hat{H}\Psi(\{\mathbf{r}\};t) = i\hbar\frac{\partial}{\partial t}\Psi(\{\mathbf{r}\};t). \quad (2.1)$$

The linear operator \hat{H} is the Hamilton operator, $\Psi(\{\mathbf{r}\};t)$ the wave function, \hbar Planck's constant divided by 2π , t the time variable and $\{\mathbf{r}\}$ the set of particle coordinates. For a closed system the Hamiltonian does not depend on t explicitly, i.e. the energy is constant in time for a given state according to the energy conservation law. These states with a defined energy E_n are called stationary states Ψ_n which are eigenfunctions of the Hamiltonian and therefore have to obey an eigenvalue equation

$$\hat{H}\Psi_n(\{\mathbf{r}\};t) = E_n\Psi_n(\{\mathbf{r}\};t). \quad (2.2)$$

A comparison with the wave equation eq. 2.1 yields the time-dependence of $\Psi_n(\{\mathbf{r}\};t)$

$$\Psi_n(\{\mathbf{r}\};t) = \psi_n(\{\mathbf{r}\})e^{-\frac{i}{\hbar}E_nt}. \quad (2.3)$$

The stationary wave functions without a time factor $\psi_n(\{\mathbf{r}\})$ and the corresponding eigenvalues are solutions to the stationary Schrödinger equation

$$\hat{H}\psi_n(\{\mathbf{r}\}) = E_n\psi_n(\{\mathbf{r}\}). \quad (2.4)$$

The Hamiltonian is of central importance in quantum mechanics since it determines the form of eq. 2.1 and eq. 2.4. For a molecular system it is

$$\begin{aligned}\hat{H} &= -\frac{1}{2} \sum_i \nabla_i^2 - \frac{1}{2} \sum_A \frac{1}{M_A} \nabla_A^2 - \sum_{i,A} \frac{Z_A}{r_{iA}} + \sum_{j<i} \frac{1}{r_{ij}} + \sum_{B<A} \frac{Z_A Z_B}{r_{AB}} \\ &= \hat{T}_e + \hat{T}_N + \hat{V}_{eN} + \hat{V}_{ee} + \hat{V}_{NN}.\end{aligned}\tag{2.5}$$

The operator is given in atomic units [30], Z_A is the nuclear charge number, M_A the nuclear mass, i, j are electronic and A, B nuclear indices. The \hat{T} denote operators of kinetic and \hat{V} of potential energies.

The two-particle operators \hat{V}_{eN} , \hat{V}_{ee} and \hat{V}_{NN} prevent an exact solution of the Schrödinger equation except for the most simple atomic system. In order to be able to treat molecules of interest approximations to the exact solution have to be introduced.

2.2 Born-Oppenheimer Approximation

Because of the large difference between nuclear and electronic masses by at least 3 orders of magnitude, the nuclei usually move much slower than the electrons. This means that the electrons can follow the changes of the nuclear frame instantaneously, which allows us to separate their motions and to describe the electron configuration in the presence of a fixed nuclear frame ($T_N = 0$, $V_{NN} = \text{const.}$). The resulting electronic Hamiltonian

$$\begin{aligned}\hat{H}_{el} &= -\frac{1}{2} \sum_i \nabla_i^2 - \sum_{i,A} \frac{Z_A}{r_{iA}} + \sum_{j<i} \frac{1}{r_{ij}} \\ &= \sum_i \hat{h}_i + \sum_{j<i} \frac{1}{r_{ij}}\end{aligned}\tag{2.6}$$

acts on the electronic coordinates and only depends parametrically on the nuclear positions. The total energy is the sum of the electronic energy E_{el} and the nuclear-nuclear repulsion energy V_{NN} . This approximation yields good results in general, but e.g. at the intersection of electronic states of same symmetry it completely fails because of the strong coupling of nuclear and electronic motions.

The Born-Oppenheimer approximation simplifies the problem to the solution of the electronic Schrödinger equation, but the spectrum of analytical solutions is only extended to problems that are isoelectronic to the hydrogen atom because of the electron-electron repulsion term \hat{V}_{ee} .

Since the following work will focus on the solution of the electronic Schrödinger equation, the index 'el' will be dropped ($\hat{H} = \hat{H}_{el}$).

2.3 Hartree-Fock Theory

In this section the most simple ab initio method to solve the electronic Schrödinger equation of a many-electron system, the Hartree-Fock method (HF), is presented. If we consider a system of non-interacting electrons (independent particle model, IPM) whose Hamiltonian is given as a sum of effective one-electron operators, the wave function can be separated and represented as an anti-symmetrized product of one-electron functions, the so called Slater determinant [31]

$$\begin{aligned} \Psi(\{\mathbf{r}\}) &= |\phi_1, \phi_2, \dots, \phi_N\rangle \\ &= (N!)^{-1} \det \begin{vmatrix} \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) & \cdots & \phi_N(\mathbf{x}_1) \\ \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) & \cdots & \phi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{x}_N) & \phi_2(\mathbf{x}_N) & \cdots & \phi_N(\mathbf{x}_N) \end{vmatrix}, \end{aligned} \quad (2.7)$$

with N as the number of electrons, ϕ_i as spin orbital i and \mathbf{x}_j as space-spin coordinates of electron j . It is easily shown that this choice for the wave function satisfies the antisymmetry principle according to Pauli [32, 33] ($\Psi(\mathbf{x}_1, \dots, \mathbf{x}_i, \mathbf{x}_j, \dots, \mathbf{x}_N) = -\Psi(\mathbf{x}_1, \dots, \mathbf{x}_j, \mathbf{x}_i, \dots, \mathbf{x}_N)$) as well as the indistinguishability of the electrons.

Resorting to the Slater-Condon rules [30] the expectation value of the Hamiltonian in eq. 2.6 with this wave function yields the HF ground-state energy

$$E_0^{HF} = \sum_i^{occ} \langle i | \hat{h} | i \rangle + \frac{1}{2} \sum_{i,j}^{occ} \langle ij || ij \rangle, \quad (2.8)$$

where the Dirac notation for the integrals is used

$$\begin{aligned} \langle i | \hat{h} | i \rangle &= \int \phi_i^*(\mathbf{x}_1) \hat{h} \phi_i(\mathbf{x}_1) d\mathbf{x}_1, \\ \langle ij || ij \rangle &= \langle ij | ij \rangle - \langle ij | ji \rangle, \\ \langle ij | ij \rangle &= \int \int \phi_i^*(\mathbf{x}_1) \phi_i(\mathbf{x}_1) r_{12}^{-1} \phi_j^*(\mathbf{x}_2) \phi_j(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2. \end{aligned} \quad (2.9)$$

According to the variational principle [30], whose equivalence to the Schrödinger equation can be easily shown [34], the energy expectation value obtained with a trial function $\tilde{\Psi}$ is

an upper bound to the true energy

$$\langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle / \langle \tilde{\Psi} | \tilde{\Psi} \rangle \geq E_0^{ex}. \quad (2.10)$$

Thus a proper trial function is chosen and optimized in such a way that the expectation value becomes minimal $\delta \langle E \rangle = 0$ with respect to arbitrary variations in $\tilde{\Psi}$. With Lagrange's method of undetermined multipliers under the constraint of orthonormal molecular orbitals ($\langle \phi_i | \phi_j \rangle = \delta_{ij}$) one obtains – after unitary transformation – the canonical HF equation [30]

$$\hat{F} \phi_i = \epsilon_i \phi_i, \quad (2.11)$$

with the Fock operator

$$\hat{F} = \hat{h} + \sum_j \left(\hat{J}_j - \hat{K}_j \right). \quad (2.12)$$

The operator \hat{J}_j corresponds to the classical Coulomb interaction, the non-classical exchange operator \hat{K}_j results directly from the antisymmetry condition

$$\hat{J}_j = \int d\mathbf{x}_2 \phi_j^*(\mathbf{x}_2) \frac{1}{r_{12}} \phi_j(\mathbf{x}_2), \quad (2.13)$$

$$\hat{K}_j = \int d\mathbf{x}_2 \phi_j^*(\mathbf{x}_2) \frac{\hat{P}_{12}}{r_{12}} \phi_j(\mathbf{x}_2), \quad (2.14)$$

where \mathbf{x}_2 is the space-spin coordinate of electron 2 and \hat{P}_{12} the permutational operator exchanging the coordinates of electrons 1 and 2. Because of the dependence of the Fock operator \hat{F} on the one-electron functions ϕ_i the HF equations have to be solved iteratively in a self-consistent field (SCF) procedure.

The molecular orbitals ϕ_i are in general represented as a linear combination of basis functions (LCBF) from a finite set of (contracted) Gaussians $\{\chi_\mu\}$

$$\phi_i = \sum_{\mu} C_{\mu i} \chi_{\mu}. \quad (2.15)$$

So the problem of orbital optimization is transformed into the algebraic problem of coefficients ($C_{\mu i}$) optimization. For a closed-shell system this leads to the Roothaan-Hall equations [35]

$$\mathbf{FC} = \mathbf{SC}\epsilon, \quad (2.16)$$

with the Fock matrix

$$F_{\mu\nu} = h_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma} \left[\langle \mu\lambda | \nu\sigma \rangle - \frac{1}{2} \langle \mu\lambda | \sigma\nu \rangle \right], \quad (2.17)$$

the metric \mathbf{S} , the molecular orbital coefficients matrix \mathbf{C} , and the diagonal matrix $\boldsymbol{\epsilon}$ containing the orbital energies. In the preceding text, matrices are denoted with bold letters. The matrix \mathbf{P} is the representation of the one-electron or Fock-Dirac density in the given basis (see the following section)

$$P_{\mu\nu} = \sum_{i \in occ} C_{\mu i} C_{\nu i}^*. \quad (2.18)$$

The energy results as

$$E_0 = \frac{1}{2} \text{Tr} [\mathbf{P} (\mathbf{h} + \mathbf{F})], \quad (2.19)$$

with \mathbf{h} as the matrix representation of the core Hamiltonian. The formal scaling behavior with the number of basis functions M of the HF method is $\mathcal{O}(M^4)$, i.e. the computational effort scales with the fourth power because of the $\frac{1}{8}M^4$ two-electron integrals (prefactor comes from the permutational symmetry of the integrals). With the introduction of direct SCF methods by Almlöf [36] and integral screening with respect to the one-particle density [37] the scaling is reduced to quadratic. New methods based on the fast multipole method (FMM) to construct the Coulomb matrix [1–4, 6, 7] and linear scaling exchange matrix formation [8, 10] enable an asymptotically linear scaling computation of the Fock-type matrices for systems with a non-vanishing HOMO-LUMO gap.

While for smaller systems the $\mathcal{O}(M^2)$ formation of the Fock matrix is clearly the most time consuming step because of its large prefactor, the cubic scaling behavior of the algebraic routines to diagonalize \mathbf{F} or to construct the density matrix becomes dominant for larger systems. A solution to this problem is given by density matrix-based SCF methods [11, 13–15, 38] that are completely formulated in the basis of local quantities and thus enable an overall linear scaling in combination with sparse algebra routines^{a)}.

^{a)}See e.g. section 3.6.1.1.

2.4 Linear Scaling Formation of the Fock Matrix

In order to analyze the scaling behavior of the formation of Coulomb and exchange matrices we have to reconsider the explicit form of the matrix elements

$$\begin{aligned} J_{\mu\nu} &= \sum_{\lambda\sigma} P_{\lambda\sigma} (\mu\nu|\lambda\sigma) \\ &= \sum_{\lambda\sigma} P_{\lambda\sigma} \int \chi_{\mu}^*(\mathbf{r}_1)\chi_{\nu}(\mathbf{r}_1)r_{12}^{-1}\chi_{\lambda}^*(\mathbf{r}_2)\chi_{\sigma}(\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2, \end{aligned} \quad (2.20)$$

$$\begin{aligned} K_{\mu\nu} &= \sum_{\lambda\sigma} P_{\lambda\sigma} (\mu\sigma|\lambda\nu) \\ &= \sum_{\lambda\sigma} P_{\lambda\sigma} \int \chi_{\mu}^*(\mathbf{r}_1)\chi_{\sigma}(\mathbf{r}_1)r_{12}^{-1}\chi_{\lambda}^*(\mathbf{r}_2)\chi_{\nu}(\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2, \end{aligned} \quad (2.21)$$

where we have used the Mulliken notation for two-electron integrals. The integrals describe the Coulomb interaction between the charge distributions for electrons 1 and 2. Since the basis functions χ_{μ} are (contracted) Gaussian functions, the number of significant function pairs $\chi_{\mu}(\mathbf{r}_i)\chi_{\nu}(\mathbf{r}_i)$ scales linearly with system size. Recalling the Gaussian product theorem the product of two Gaussians yields a new Gaussian centered on the line joining the original functions which is scaled by the exponential $\exp(-C R_{12}^2)$, i.e. the product decays exponentially with the distance R_{12} between the original Gaussian. Therefore the number of significant charge distributions scales asymptotically linear with system size and the formation of two-electron integrals with $\mathcal{O}(M^2)$.

To achieve an $\mathcal{O}(M)$ behavior the Coulomb and exchange matrices have to be treated separately. Since the bra and ket functions of the Coulomb integral in eq. 2.20 are coupled by r_{12}^{-1} only, the number of integrals scales quadratically with M . This can be reduced by the application of the continuous fast multipole method (CFMM) introduced by White et al. [1–4] who extended the FMM method of Greengard and Rokhlin [39] to the treatment of extended charge distributions. Note that there are similar approaches proposed by several authors [6, 7], but the new developments presented in this thesis are based on the CFMM method which is the standard method used in the Q-CHEM package [40]. The Coulomb matrix is split into near-field and far-field contributions with respect to the distance of the basis function pairs $(\chi_{\mu}\chi_{\nu})$ and $(\chi_{\lambda}\chi_{\sigma})$, where the near-field part is computed with standard integral routines and the far-field is taken into account as an interaction with a multipole field arising from distant charge distributions.

The number of significant elements in the exchange matrix, on the other hand, scales

linearly with M for a system with a non-vanishing HOMO-LUMO gap because of the coupling of the bra- and ket-sides by the density matrix \mathbf{P} . As will be shown in the following section, the density matrix has only a linear scaling number of significant elements for systems with a local electronic structure. Thus the bra- ($\chi_\mu\chi_\sigma$) and ket-terms ($\chi_\lambda\chi_\nu$) are coupled by a constant number of density matrix elements $P_{\lambda\sigma}$. This means that each bra-pair ($\chi_\mu\chi_\sigma$) has only a constant number of significant $P_{\lambda\sigma}$ which again has a constant number of ket-pairs ($\chi_\lambda\chi_\nu$). The problem in forming the exchange matrix in eq. 2.21 emerges from the quadratically scaling screening scheme, which would asymptotically dominate the exchange matrix formation. This obstacle was first removed by the order N exchange (ONX) method proposed by Schwegler et al. [10]. However, their method does not exploit the permutational symmetry of the two-electron integrals and therefore introduced a prefactor which becomes crucial for systems with limited extend or small band gaps. Thus, Ochsenfeld et al. introduced the LinK scheme [8] which is also used throughout this thesis. Since the LinK screening has only a small overhead compared to the standard routines it is competitive even for small systems.

2.5 The Electron Density

Since the methods that have been developed in this work are all based on reduced electron density matrices, their properties are discussed in this section.

In the framework of quantum mechanics the complete description of a system is given by the wave function $\Psi(\{x\}, t)$. The Hartree-Fock and the Kohn-Sham reference wave functions, which are in center of the following discussion, are represented by an antisymmetrized product of orthonormal one-electron functions $\{\phi_i\}$ (spin orbitals), i.e. the Slater determinant introduced earlier (eq. 2.7). The invariance of a single determinant to a unitary transformation of the spin orbitals [30] shows, that the spin orbitals do not represent a unique set of solutions to the operator in eq. 2.12. Because of this arbitrariness and the fact, that canonical molecular orbitals are normally delocalized over the complete molecule, the MOs do not provide a unique physical picture of the electronic structure of the system.

The information about the distribution of electrons as well as clusters of electrons is given by the n -particle reduced densities ρ_n where n is the number of electrons in the

corresponding cluster. Following the interpretation according to Born [33],

$$|\Psi|^2 d\mathbf{x}_1 d\mathbf{x}_2 \dots d\mathbf{x}_N = \Psi^*(\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_N) \Psi(\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_N) d\mathbf{x}_1 d\mathbf{x}_2 \dots d\mathbf{x}_N \quad (2.22)$$

is the probability for the configuration with electron 1 in $d\mathbf{x}_1$, electron 2 in $d\mathbf{x}_2$ etc. The probability to find an arbitrary electron at \mathbf{x}_1 independently of the positions of the remaining electrons equals – because of the indistinguishability of fermions – the probability of electron 1 at \mathbf{x}_1 multiplied by N . This defines the 1-particle reduced density function

$$\rho_1(\mathbf{x}_1) = N \int \Psi(\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_N) \Psi^*(\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_N) d\mathbf{x}_2 \dots d\mathbf{x}_N. \quad (2.23)$$

In the context of density functions the index of the variable \mathbf{x}_1 does not refer to electron 1 but to a point where the density is evaluated. Note that these functions are – just like an observable – quadratic in Ψ and so invariant to unitary transformations of the spin orbitals. A further important common property with 1-particle operators is the time-dependency (see sec. 2.5.2).

In the following the 1- and 2-particle expectation values of the Hamiltonian in eq. 2.5 are represented as functions of the 1- and 2-particle reduced densities, respectively. This is obviously no problem for purely multiplicative operators, but the differential operator of the kinetic energy requires a more general form of the electron density. We define the 1-particle reduced density matrix as

$$\rho_1(\mathbf{x}_1; \mathbf{x}'_1) = N \int \Psi(\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_N) \Psi^*(\mathbf{x}'_1 \mathbf{x}_2 \dots \mathbf{x}_N) d\mathbf{x}_2 \dots d\mathbf{x}_N, \quad (2.24)$$

so that it is possible to define the expectation value of any 1-particle operator \hat{O}_1 as

$$\langle \hat{O}_1 \rangle = \int_{\mathbf{x}'_1 = \mathbf{x}_1} \hat{O}_1 \rho_1(\mathbf{x}_1; \mathbf{x}'_1) d\mathbf{x}_1. \quad (2.25)$$

This equation has to be read as follows: The operator acts only on the coordinate \mathbf{x}_1 while \mathbf{x}'_1 is "inert", then one replaces \mathbf{x}'_1 with \mathbf{x}_1 and finally integrates over \mathbf{x}_1 . Eq. 2.24 defines a real matrix with continuous indices and the density function in eq. 2.23 as diagonal elements.

In an analogous way to the electron density one defines the electron pair distribution^{b)}

$$\rho_2(\mathbf{x}_1, \mathbf{x}_2) = N(N-1) \int \Psi(\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_N) \Psi^*(\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_N) d\mathbf{x}_3 \dots d\mathbf{x}_N, \quad (2.26)$$

^{b)}The prefactor $N(N-1)$ is due to the number of electron pairs.

which gives the probability of one electron at \mathbf{x}_1 and another at \mathbf{x}_2 . This function contains the complete information about (dynamic) correlation and describes the correlated motion of two electrons resulting from their direct interaction (see sec. 2.6).

Since the Hamiltonian in eq. 2.5 consists of 1- and 2-particle operators only, the exact (non-relativistic) energy can be written as

$$\begin{aligned}
 E_0 &= -\frac{1}{2} \int \nabla^2 \rho(\mathbf{x}_1; \mathbf{x}'_1) d\mathbf{x}_1 + \sum_A \int \frac{Z_A}{|\mathbf{r}_A - \mathbf{x}_1|} \rho(\mathbf{x}_1) d\mathbf{x}_1 \\
 &+ \frac{1}{2} \int \mathbf{r}_{12}^{-1} \pi(\mathbf{x}_1, \mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2.
 \end{aligned}
 \tag{2.27}$$

Since no other density functions are needed to describe a Coulombic system, we drop the index for the 1-electron density ($\rho = \rho_1$) and abbreviate the pair density as ($\pi = \rho_2$) in eq. 2.27.

2.5.1 1-particle Reduced Density Function of a Single Slater Determinant: Fock-Dirac Density

After the more general definitions in the previous section the form of the electron density constructed from a single determinant is derived. Instead of calculating the density from eq. 2.23, we extract it from the Hartree-Fock energy expression in eq. 2.8. The expectation value of the core Hamiltonian \hat{h} is given by

$$\begin{aligned}
 \langle \hat{h} \rangle &= \sum_{i \in occ} \langle i | \hat{h} | i \rangle \\
 &= \sum_{i \in occ} \int \phi_i^*(\mathbf{x}) \hat{h} \phi_i(\mathbf{x}) d\mathbf{x}
 \end{aligned}
 \tag{2.28}$$

or, alternatively, in density form by

$$\langle \hat{h} \rangle = \int_{\mathbf{x}'=\mathbf{x}} \hat{h} \rho(\mathbf{x}; \mathbf{x}') d\mathbf{x}.
 \tag{2.29}$$

In a similar way to the expansion of an arbitrary wave function in the basis of eigenstates, one can represent the density for an arbitrary pure state as a linear combination of the densities of the stationary states [32, 41]. In the case of an IPM $\rho(\mathbf{x}; \mathbf{x}')$ is expanded in the basis of one-electron molecular spin-orbital products

$$\rho(\mathbf{x}; \mathbf{x}') = \sum_{rs} p_{rs} \phi_r(\mathbf{x}) \phi_s^*(\mathbf{x}'),
 \tag{2.30}$$

where the coefficients p_{rs} form the matrix representation \mathbf{P}^{MO} of an abstract density operator in the basis $\{\phi_i\}$. A comparison with eq. 2.28 and eq. 2.29 shows that \mathbf{P}^{MO} has a diagonal form where the diagonal elements are the occupation numbers of the one-electron orbitals

$$p_{rs} = \begin{cases} \delta_{rs} & \text{if } r, s \in occ \\ 0 & \text{else} \end{cases}. \quad (2.31)$$

This leads to the so called Fock-Dirac density

$$\rho(\mathbf{x}; \mathbf{x}') = \sum_{i \in occ} \phi_i(\mathbf{x}) \phi_i^*(\mathbf{x}'). \quad (2.32)$$

The corresponding density operator $\hat{\rho}$ is the superposition of all MO dyadics of the occupied subspace

$$\hat{\rho} = \sum_{i \in occ} |\phi_i\rangle \langle \phi_i|. \quad (2.33)$$

From the repulsive 2-electron term in eq. 2.8 one finds for the pair density

$$\begin{aligned} V_{ee}^{HF} &= \frac{1}{2} \sum_{i,j}^{occ} \langle ij | ij \rangle - \langle ij | ji \rangle \\ &= \frac{1}{2} \int_{\substack{\mathbf{x}'_1 = \mathbf{x}_1 \\ \mathbf{x}'_2 = \mathbf{x}_2}} \mathbf{r}_{12}^{-1} \pi(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) d\mathbf{x}_1 d\mathbf{x}_2, \\ \implies \pi(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) &= \rho(\mathbf{x}_1; \mathbf{x}'_1) \rho(\mathbf{x}_2; \mathbf{x}'_2) - \rho(\mathbf{x}_2; \mathbf{x}'_1) \rho(\mathbf{x}_1; \mathbf{x}'_2). \end{aligned} \quad (2.34)$$

This is a special case of the general result, that each n -particle reduced density matrix ρ_n can be expressed as an antisymmetrized product of Fock-Dirac density matrices, i.e. by a Laplace expansion of the Slater-determinant in the first n rows [42, 43]

$$\rho_n(\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_n; \mathbf{x}'_1 \mathbf{x}'_2 \dots \mathbf{x}'_n) = \begin{vmatrix} \rho_1(\mathbf{x}_1; \mathbf{x}'_1) & \rho_1(\mathbf{x}_1; \mathbf{x}'_2) & \cdots & \rho_1(\mathbf{x}_1; \mathbf{x}'_n) \\ \rho_1(\mathbf{x}_2; \mathbf{x}'_1) & \rho_1(\mathbf{x}_2; \mathbf{x}'_2) & \cdots & \rho_1(\mathbf{x}_2; \mathbf{x}'_n) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_1(\mathbf{x}_n; \mathbf{x}'_1) & \rho_1(\mathbf{x}_n; \mathbf{x}'_2) & \cdots & \rho_1(\mathbf{x}_n; \mathbf{x}'_n) \end{vmatrix}. \quad (2.35)$$

This expansion is another sign for the independent-particle character of the approximation, which describes the N -electron wave function as antisymmetrized product of N one-electron functions, and states the complete determination of the system by the Fock-Dirac density.

The treatment of the density so far has been in the continuous basis $\{\mathbf{r}\}$ of 3-dimensional space. Since the methods that are in the focus of this work are in general formulated within a finite set of basis functions, a discrete description has to be introduced. From the expansion of the molecular orbitals in a basis $\{\chi_\mu\}$ (eq. 2.15) one gets the corresponding matrix representation \mathbf{P} of the density by transformation from the MO basis $\{\phi_i\}$ to the atomic orbital (AO) basis $\{\chi_\mu\}$

$$\begin{aligned} \mathbf{P} &= \mathbf{C}\mathbf{P}^{MO}\mathbf{C}^\dagger, \\ \Rightarrow P_{\mu\nu} &= \sum_{i \in occ} C_{\mu i} C_{\nu i}^*. \end{aligned} \quad (2.36)$$

So far the spin-dependency was implicitly included by the use of spin-orbitals. Introducing α - and β -spin densities one obtains

$$\begin{aligned} P_{\mu\nu} &= P_{\mu\nu}^\alpha + P_{\mu\nu}^\beta, \\ P_{\mu\nu}^\alpha &= \sum_{i \in occ_\alpha} C_{\mu i}^\alpha C_{\nu i}^{\alpha*}, \\ P_{\mu\nu}^\beta &= \sum_{i \in occ_\beta} C_{\mu i}^\beta C_{\nu i}^{\beta*}. \end{aligned} \quad (2.37)$$

2.5.2 Properties of the Fock-Dirac Density

One special property can immediately be seen by applying the density operator to an arbitrary function f that is expanded in the basis $\{\phi_r\}$

$$f = \sum_r a_r |\phi_r\rangle. \quad (2.38)$$

Due to the orthonormality of the spin-orbitals, $\hat{\rho}$ projects the function onto the occupied subspace

$$\begin{aligned} \hat{\rho}f &= \sum_{i \in occ} \sum_r a_r |\phi_i\rangle \langle \phi_i | \phi_r\rangle = \sum_{i \in occ} \sum_r a_r |\phi_i\rangle \delta_{ir} \\ &= \sum_{i \in occ} a_i |\phi_i\rangle. \end{aligned} \quad (2.39)$$

This projection onto the occupied subspace is essential for the upcoming sections. The idempotency, which is a necessary condition for a projector, results directly from the or-

thonormality of the molecular orbitals

$$\begin{aligned}
 \hat{\rho}^2 &= \sum_{\substack{i \in occ \\ j \in occ}} |\phi_i\rangle \langle \phi_i | \phi_j\rangle \langle \phi_j| = \sum_{\substack{i \in occ \\ j \in occ}} |\phi_i\rangle \langle \phi_j| \delta_{ij} \\
 &= \sum_{i \in occ} |\phi_i\rangle \langle \phi_i| = \hat{\rho}.
 \end{aligned} \tag{2.40}$$

Within a discrete representation of the density we have to consider that the basis functions are in general normalized but non-orthogonal. Thus the metric $S_{\mu\nu} = \langle \chi_\mu | \chi_\nu \rangle$ of the space spanned by the M functions χ_μ has to be considered explicitly. With

$$\langle \phi_r | \phi_s \rangle = \delta_{rs} \implies \sum_{\mu\nu} C_{\mu r}^* \langle \chi_\mu | \chi_\nu \rangle C_{\nu s} = \sum_{\mu\nu} C_{\mu r}^* S_{\mu\nu} C_{\nu s} = \delta_{rs}, \tag{2.41}$$

we obtain the idempotency relation in a non-orthogonal basis

$$\mathbf{PSP} = \mathbf{P}. \tag{2.42}$$

Furthermore, it can be seen from eq. 2.41 that the trace of the Fock-Dirac density equals the number of electrons [42, 44]

$$\text{Tr}[\mathbf{PS}] = N, \tag{2.43}$$

since the trace is invariant with respect to cyclic permutations of its arguments

$$\begin{aligned}
 \text{Tr}[\mathbf{PS}] &= \text{Tr}[\mathbf{CP}^{MO}\mathbf{C}^\dagger\mathbf{S}] = \text{Tr}[\mathbf{P}^{MO}\mathbf{C}^\dagger\mathbf{SC}] \\
 &= \text{Tr}[\mathbf{P}^{MO}\mathbf{1}] = N.
 \end{aligned} \tag{2.44}$$

Finally, as the last property, the time-dependency of the electron density can be derived from comparison of the time derivative of the density with the wave equation in eq. 2.1 [45], giving a von-Neumann type equation of motion (in a.u.)

$$\begin{aligned}
 \frac{\partial}{\partial t} \hat{\rho} &= |\dot{\Psi}\rangle \langle \Psi^*| + |\Psi\rangle \langle \dot{\Psi}^*| \\
 &= -i\hat{H}|\Psi\rangle \langle \Psi^*| + i|\Psi\rangle \langle \Psi^*|\hat{H} \\
 &= -i\left(\hat{H}\hat{\rho} - \hat{\rho}\hat{H}\right), \\
 \implies i\mathbf{S}\dot{\mathbf{P}}\mathbf{S} &= \mathbf{FPS} - \mathbf{SPF}
 \end{aligned} \tag{2.45}$$

with $\dot{\mathbf{P}} = \partial\mathbf{P}/\partial t$ and $\dot{\Psi} = \partial\Psi/\partial t$. Since the density of a stationary state is time-independent, it behaves like a constant of motion and commutes with the Hamiltonian

$$\mathbf{0} = \mathbf{FPS} - \mathbf{SPF}. \tag{2.46}$$

From this equation and eq. 2.36 it is easy to see that \mathbf{P} has to be self-adjoint

$$\mathbf{P}^\dagger = \mathbf{P}. \quad (2.47)$$

As mentioned before the density is the projector onto the occupied subspace. Including the non-orthogonality of the basis, we define the projectors onto the occupied ($\mathbf{P}_{\text{occ}} = \mathbf{P}\mathbf{S}$) and virtual ($\mathbf{P}_{\text{virt}} = (\mathbf{1} - \mathbf{P}\mathbf{S})$) subspaces. The following relations between \mathbf{P}_{occ} and \mathbf{P}_{virt} hold:

$$\mathbf{P}_{\text{occ}}\mathbf{P}_{\text{virt}} = \mathbf{P}_{\text{virt}}\mathbf{P}_{\text{occ}} = \mathbf{0}, \quad (2.48)$$

$$\mathbf{P}_{\text{occ}} + \mathbf{P}_{\text{virt}} = \mathbf{1}. \quad (2.49)$$

The first equation shows the orthogonality of the occupied and virtual subspaces, the second results from the resolution of unity since it equals the superposition of all MO dyadics. The latter can be used to split any matrix representation \mathbf{A} of a general operator \hat{A} in the given basis into a sum of subspace projections^{c)}:

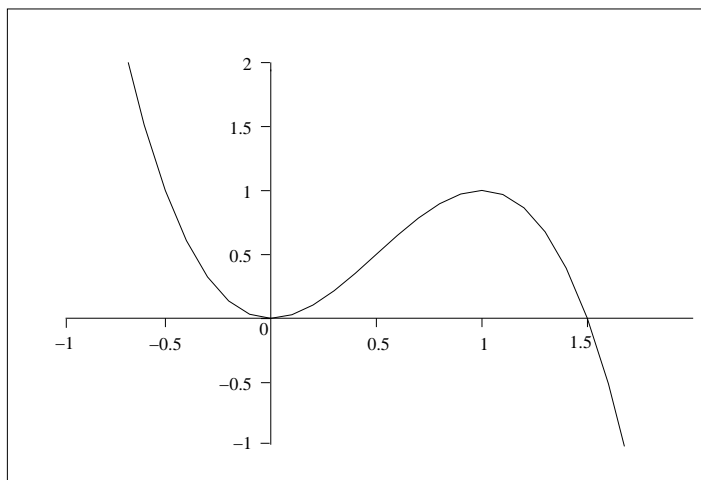
$$\begin{aligned} \mathbf{A} &= \mathbf{1}\mathbf{A}\mathbf{1} \\ &= \mathbf{S}\mathbf{P}\mathbf{A}\mathbf{S} + \mathbf{S}\mathbf{P}\mathbf{A}(\mathbf{1} - \mathbf{P}\mathbf{S}) + (\mathbf{1} - \mathbf{S}\mathbf{P})\mathbf{A}\mathbf{S} + (\mathbf{1} - \mathbf{S}\mathbf{P})\mathbf{A}(\mathbf{1} - \mathbf{P}\mathbf{S}) \\ &= \mathbf{A}_{\text{oo}} + \mathbf{A}_{\text{ov}} + \mathbf{A}_{\text{vo}} + \mathbf{A}_{\text{vv}}. \end{aligned} \quad (2.50)$$

This unique resolution of matrix representations of arbitrary operators into their subspace projections will play a central role in the upcoming derivations of D-CPSCF/TDSCF equations.

2.5.2.1 Purification Transformation

Let us reflect on the four properties of the density matrix in a standard SCF treatment. In the optimization process the density matrices are built from the MO coefficient vectors \mathbf{C} that result from the diagonalization of the Fock or Kohn-Sham matrices. It is obvious that the self-adjointness, N -trace as well as the idempotency are satisfied after the diagonalization even for a non-converged density matrix. The commutator in eq. 2.46 holds only for the converged density and so its norm is a standard test for convergence.

^{c)}Assuming a covariant matrix \mathbf{A} . For an introduction to tensor theory in many-electron theory see e.g. [46].

Figure 2.1: "Purification" function $\tilde{x} = 3x^2 - 2x^3$ 

In an overall linear scaling SCF algorithm the explicit use of the MO coefficient vectors is prohibitive because of the dense structure of \mathbf{C} , so the goal has to be the direct solution for the density matrix. The self-adjointness of \mathbf{P} will directly emerge from the symmetry of the corresponding linear equation systems which will be presented in the following sections (see e.g. eq. 3.72). A more involved problem is the enforcement of the idempotency in the optimization process (wherefrom the N -trace property emerges if the initial density represents the correct number of electrons). Ignoring these constraints in the minimization, the eigenvalues of the density would tend to $-\infty$ for the occupied and to $+\infty$ for the virtual eigenstates.

The key tool to impose these constraints – at least to first order – in the methods presented in the following is the purification transformation introduced by McWeeny [47, 48] (see eq. 2.51 and Fig. 2.1).

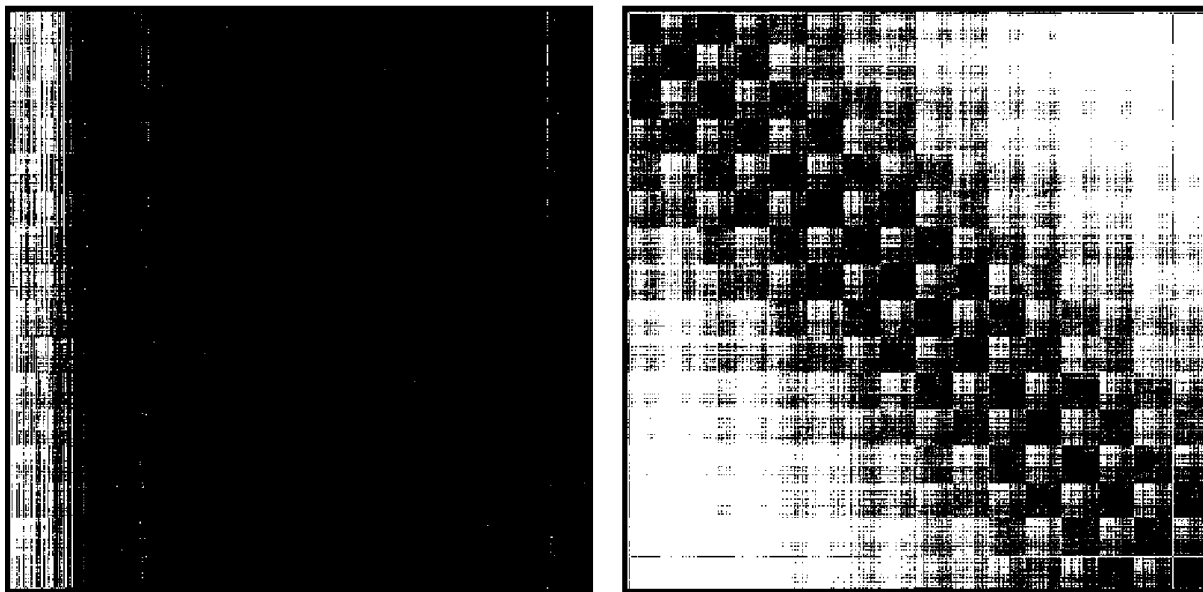
A nearly idempotent density matrix \mathbf{P} has eigenvalues n_r (\rightarrow occupation numbers) close to 0 or 1, i.e. $1 - \theta$ or θ with $0 \leq \theta \ll 1$. Inserting this trial density into

$$\tilde{\mathbf{P}} = 3\mathbf{P}\mathbf{S}\mathbf{P} - 2\mathbf{P}\mathbf{S}\mathbf{P}\mathbf{S}\mathbf{P}, \quad (2.51)$$

we obtain a more nearly idempotent density matrix $\tilde{\mathbf{P}}$ as can easily be seen from the function plot in Fig. 2.1:

$$\begin{aligned} f(1 - \theta) &\rightarrow 1 \geq (1 - 3\theta^2 + 2\theta^3) \geq 1 - \theta, \\ f(\theta) &\rightarrow 0 \leq (3\theta^2 - 2\theta^3) \leq \theta. \end{aligned}$$

Figure 2.2: Sparsity pattern of the MO coefficient matrix (left) and the one-electron density (right) of a DNA fragment containing 8 base pairs from a HF/6-31G* calculation. The white dots are non-significant elements with an absolute value smaller than 10^{-7} .



Note that this transformation only works for values in the range $[-0.5, 1.5]$ and the equalities hold for $n_r = 0 \vee 1$ ($\theta = 0$) [47, 48]. Within an optimization procedure we are now able to impose the idempotency condition either by inserting $\tilde{\mathbf{P}}$ into the energy functional (see section 3.6.1.1) or by a subsequent application of the purification transformation. As can be seen in Fig. B.1 in the appendix, the idempotency of the density is restored within a small number of iterations.

2.5.3 Structure of the Fock-Dirac Density

As mentioned before the molecular orbitals are in general delocalized over the whole nuclear frame, i.e. the MO coefficient matrix is dense as shown on the left of Fig. 2.2. As has been shown (see e.g. Refs. [49, 50]), the number of significant elements of the density matrix \mathbf{P} , whose pattern is shown on the right of Fig. 2.2, scales linearly with system size for molecules with a non-vanishing HOMO-LUMO gap.

With Born's interpretation of $|\Psi|^2$ in mind we can interpret the off-diagonal elements of the 1-particle reduced density matrix in eq. 2.24 as a measure of the localization/delocalization

of the electrons. While the diagonal elements ($x' = x$) give the probability of finding an electron at x , the off-diagonal elements give the probability of finding a delocalized electron simultaneously at x and $x' = x + \Delta x$. The connection between the HOMO-LUMO gap and the local or non-local nature of the electron structure, i.e. the scaling behavior of the number of significant elements of \mathbf{P} , can be seen for the example of a Peierls distortion of a hydrogen chain. As shown in Fig. B.2 in the appendix, the metallic system with equidistant protons does not show a HOMO-LUMO gap, i.e. it is a conductive system with electrons delocalized over the whole chain (broken line). On the other hand the chain of hydrogen molecules shows a significant gap and thus a localized electron structure (full line) with a delocalization radius Δx of approx. 44 Å instead of half of the chain length ($\Delta x \approx 167$ Å), considering the center of the chain as reference point x (threshold $< 10^{-7}$).

2.6 Electron Correlation

Typically the Hartree-Fock energy accounts for approx. 99% of the non-relativistic energy. The error results from the assumption of independent particles that only interact indirectly over an averaged field. This means, that the motions of two electrons are not correlated, i.e. the probability of finding electron 2 at r_2 does not depend on the position r_1 of electron 1. Note that the exchange term in HF theory introduces a partial correlation between electrons with parallel spins while electron pairs with opposite spins remain completely uncorrelated.

The difference between the exact, non-relativistic energy E_{ex} and the HF energy obtained with a complete basis E_{HF}^{limit} is called the correlation energy

$$E_{corr} = E_{ex} - E_{HF}^{limit}. \quad (2.52)$$

It is noticeable that the correlation energy is not a real physical quantity, i.e. it is not an observable and only an artefact of the independent particle approximation. Eq. 2.52 is a pragmatic definition of the electron correlation energy since most methods are based on the HF reference. In the following text we will also refer to the Fermi-correlation which arises from the exchange term in eq. 2.14, even if it is excluded by the definition in eq. 2.52.

The correlation energy E_{corr} contributes only approx. 1% to the total energy, but it has to be considered in many cases since it is clearly on the order of chemical energies. Compared to the HF calculation, the determination of the correlation energy by Post-HF

methods requires huge computational effort because of the unfavorable scaling behavior (see Fig. 2.4).

2.6.1 Pair Functions in HF-Theory

As mentioned in sec. 2.5 the pair density $\pi(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2)$ contains information on the correlated motions of two electrons. The HF pair density is given in eq. 2.34 as

$$\pi(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) = \rho(\mathbf{x}_1; \mathbf{x}'_1)\rho(\mathbf{x}_2; \mathbf{x}'_2) - \rho(\mathbf{x}_2; \mathbf{x}'_1)\rho(\mathbf{x}_1; \mathbf{x}'_2). \quad (2.53)$$

This means that within the single determinant approximation the complete system is described by the Fock-Dirac density. A further decomposition by integrating over the spin variables yields

$$\begin{aligned} \pi(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) &= \pi^{\alpha\alpha; \alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}'_1, \mathbf{r}'_2) + \pi^{\alpha\beta; \alpha\beta}(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}'_1, \mathbf{r}'_2) \\ &+ \pi^{\beta\alpha; \beta\alpha}(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}'_1, \mathbf{r}'_2) + \pi^{\beta\beta; \beta\beta}(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}'_1, \mathbf{r}'_2) \end{aligned} \quad (2.54)$$

and a comparison with eq. 2.53 gives the the diagonal elements of the pair densities with same ($\omega_1 = \omega_2$) and opposite ($\omega_1 \neq \omega_2$) spin

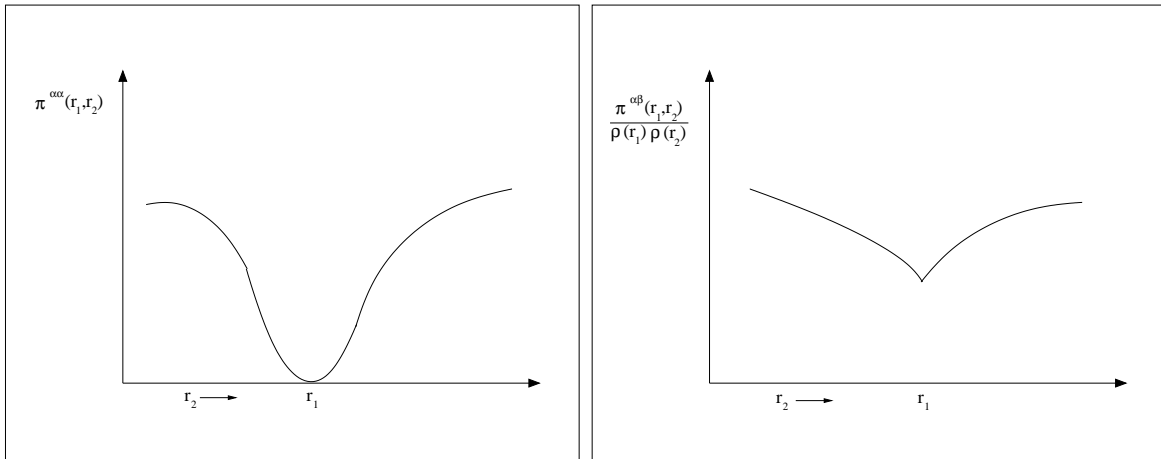
$$\pi^{\alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2) = \rho^\alpha(\mathbf{r}_1)\rho^\alpha(\mathbf{r}_2) - \rho^\alpha(\mathbf{r}_2; \mathbf{r}_1)\rho^\alpha(\mathbf{r}_1; \mathbf{r}_2) \quad (2.55)$$

$$\pi^{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_2) = \rho^\alpha(\mathbf{r}_1)\rho^\beta(\mathbf{r}_2) \quad (2.56)$$

and equivalent equations for $\pi^{\beta\beta}$ and $\pi^{\beta\alpha}$. These equations show how electron correlation is considered in the framework of HF theory. Eq. 2.55 shows that electron pairs of same spin are described by a function which vanishes for $\mathbf{r}_2 = \mathbf{r}_1$

$$\lim_{\mathbf{r}_2 \rightarrow \mathbf{r}_1} \rho^\alpha(\mathbf{r}_2; \mathbf{r}_1) = \rho^\alpha(\mathbf{r}_1) \rightarrow \lim_{\mathbf{r}_2 \rightarrow \mathbf{r}_1} \pi^{\alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2) = 0. \quad (2.57)$$

This case is depicted on the left in Fig. 2.3 and results directly from the Pauli principle that requires an antisymmetric wave function. The pair density of two electrons of opposite spins eq. 2.56 in HF theory is just the product of the corresponding one-particle densities, i.e. the motion of electron 2 would be independent of the position of electron 1. In a full interacting system the probability of finding electron 2 close to \mathbf{r}_1 is reduced as shown on the right in Fig. 2.3. This effect, which is completely neglected in HF theory, results from the Coulomb correlation emerging from the electrostatic repulsion between two charged particles.

Figure 2.3: Fermi hole (left), exact Coulomb hole (right) [34]

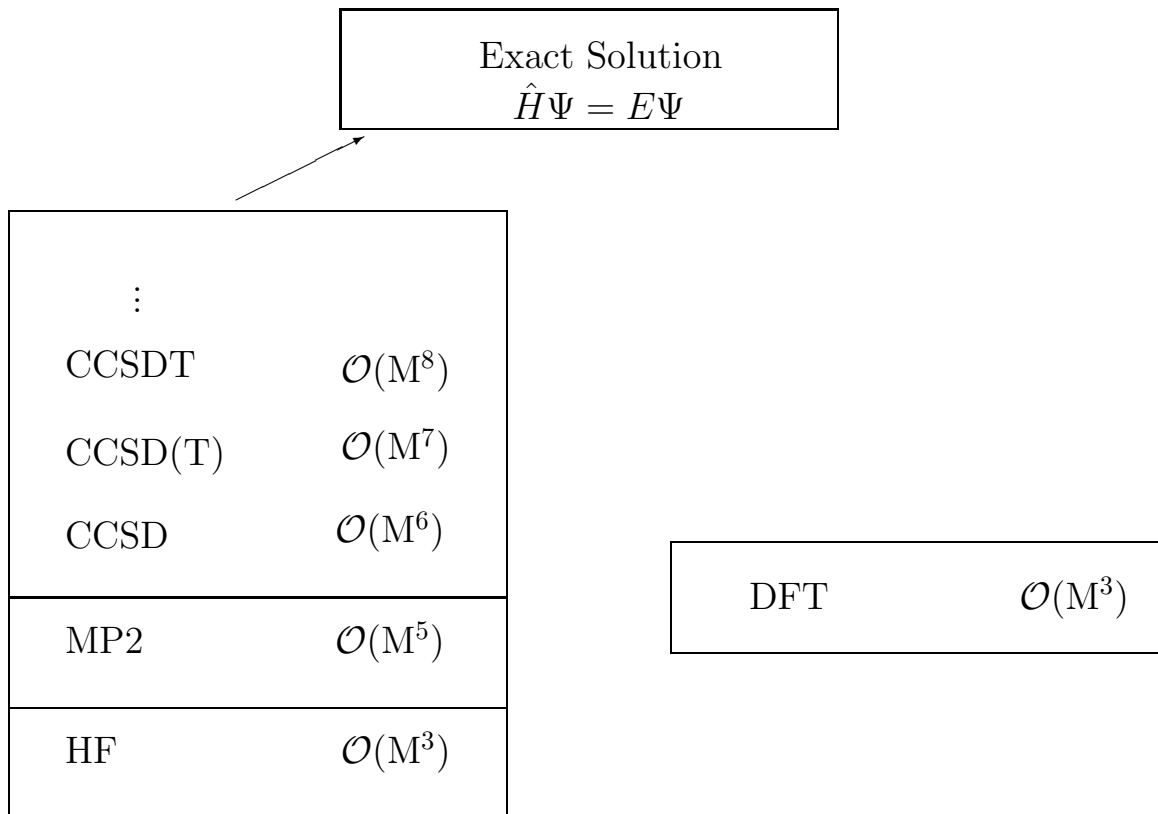
In general two types of correlation are distinguished. Dynamic correlation, which has been discussed so far, is of course local by nature and occurs in many-electron systems. This behavior of the pair density on the right in Fig. 2.3 can also be considered as consequence of the so called cusp conditions which have been formulated by Kato [51]. The discontinuity at \mathbf{r}_1 ensures an overall finite energy by a compensation of the divergence in the potential energy by a corresponding divergence in the kinetic energy^{d)}. In configuration space the dynamic correlation is introduced by a superposition of a large number of "excited" determinants (see eq. 2.58).

The static correlation, on the other side, has a multi-center character and is represented by a linear combination of near-degenerate determinants. The classic example is the dissociation of the hydrogen molecule, where a mixing of the singlet states σ_g and σ_u enables a correct description of the dissociation.

2.6.2 Expansion of a Many-Electron Wave Function in the Basis of One-Electron Functions: Full-CI Wave Functions

According to the expansion of one-electron functions in a complete basis $\{\chi_\mu\}$, one can also represent an arbitrary antisymmetric many-electron function in this basis. The exact wave function is a linear combination in the basis of all N -particle determinants (full

^{d)}The cusp conditions are also in the focus of interest when constructing correlated wave functions for QMC (see chapter 5)

Figure 2.4: Hierarchy of correlation methods.

configuration interaction, FCI)

$$\Phi_0 = c_0\Psi_0 + \sum_{ia} c_i^a\Psi_i^a + \sum_{\substack{i<j \\ a<b}} c_{ij}^{ab}\Psi_{ij}^{ab} + \sum_{\substack{i<j<k \\ a<b<c}} c_{ijk}^{abc}\Psi_{ijk}^{abc} + \dots \quad (2.58)$$

The diagonalization of the Hamiltonian in the basis of the N -electron functions in eq. 2.58 would yield the exact, non-relativistic energy of the system. However, the computational effort makes the computation not feasible for most molecules of interest due to the large number of determinants. Applying an intermediate normalization ($c_0 = 1$, $\langle\Psi_0|\Phi_0\rangle = 1$) we can derive an expression for the correlation energy

$$E_{corr} = \sum_{\substack{i<j \\ a<b}} c_{ij}^{ab}\langle\Psi_0|\hat{H}|\Psi_{ij}^{ab}\rangle. \quad (2.59)$$

Even though this equation contains only matrix elements coupling the ground state wave function and "double-excited" determinants, its solution remains a difficult task. Behind

this simple equation stands a hierarchy of coupled systems of equations that also contain higher-excited determinants. Thus one has to solve the complete hierarchy of equations to obtain a rigorous solution for c_{ij}^{ab} .

Since the description of electron correlation is a challenging and important task of quantum chemistry, a great variety of approximations has been developed. These methods like many-body perturbation theory (MBPT) [52], truncated CI or coupled cluster theory (CC) [53–55] are – although strongly different in the statement – all approximations to determine eq. 2.59 by decoupling the equation systems [56–58]. They all form a hierarchy of methods (Fig. 2.4) which provides a systematic way to the exact FCI result.

Apart from the "Post-HF" methods there exist many other approaches to seize correlation effects. The most popular alternative, the density functional theory (DFT), is presented in the following section. In the last years there was also an increasing interest in quantum Monte Carlo methods (QMC) which solve the Schrödinger equation by stochastic methods. The foundations of these methods will be summarized in the last chapter of this work where a new approach to the variational and diffusion quantum Monte Carlo schemes is presented.

2.6.3 Density Functional Theory

While Kohn-Sham DFT has been a well-established method in solid-state physics, this method was introduced to the computational chemistry community by a reformulation within a finite basis set [59–63]. Nowadays all popular *ab initio* packages provide a variety of exchange-correlation (XC) functionals that are widely used in computational chemistry and physics.

In 1927 Thomas and Fermi formulated a model for describing atoms based on a statistical approach of the uniform electron gas. In this quite simple theory the basic quantity is the electron density instead of a wave function. Nevertheless, since it can not be applied to molecular systems (Teller non-binding theorem [64]) and the results for atoms did not reach the accuracy of other methods, the impact of the Thomas-Fermi model in solid-state and molecular physics was rather small.

With the seminal work of Hohenberg and Kohn in 1964 [65] the Thomas-Fermi model was reconsidered as an approximation to a formally exact theory, the density functional theory (DFT). In the following text the foundation of this theory, the Hohenberg-Kohn

theorems, and its established manifestation in form of the Kohn-Sham DFT method are briefly discussed.

2.6.3.1 Hohenberg-Kohn Theorems

The ground state of a N -electron system is completely determined by the external potential v_{ext} in which the electrons move, i.e. the attractive potential of the nuclear frame and applied electromagnetic fields for example. The first theorem of Hohenberg and Kohn states that the electron density ρ with $\int \rho(\mathbf{r})d\mathbf{r} = N$ also provides a complete description of the system: The electron density ρ determines the external potential v_{ext} . This theorem is easily proven by contradiction. Consider for an N -electron system two external potentials v_{ext} and v'_{ext} that differ by more than an arbitrary constant and provide the same density ρ . The Hamiltonians and wave functions to the single potentials are \hat{H} , Ψ and \hat{H}' , Ψ' , respectively. Using Ψ'_0 as trial function for the Hamiltonian \hat{H} and vice versa yields

$$E_0 < \langle \Psi' | \hat{H} | \Psi' \rangle = E'_0 + \int \rho(\mathbf{r})[v_{ext}(\mathbf{r}) - v'_{ext}(\mathbf{r})], \quad (2.60)$$

$$E'_0 < \langle \Psi | \hat{H}' | \Psi \rangle = E_0 - \int \rho(\mathbf{r})[v_{ext}(\mathbf{r}) - v'_{ext}(\mathbf{r})], \quad (2.61)$$

since the the Hamiltonians only differ in the external potential. The sum of these inequalities is a contradiction $E_0 + E'_0 < E_0 + E'_0$ which shows that there is only one unique v_{ext} for a given ρ .

The second HK theorem provides a variational principle for ρ : A trial density $\tilde{\rho}$ with $\tilde{\rho} \geq 0$ and $\int \rho(\mathbf{r})d\mathbf{r} = N$ gives an energy value $E_0(\tilde{\rho})$ which is always an upper bound to the true ground state energy. Since the trial density determines its own potential \tilde{v}_{ext} and wave function $\tilde{\Psi}$, the equality to the traditional variation principle is easily seen

$$\langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle = \tilde{E}_0(\tilde{\rho}) \geq E_0. \quad (2.62)$$

With this at hand Hohenberg and Kohn were able to calculate the ground state energy in terms of the density ρ

$$\begin{aligned} E(\rho) &= \int \rho(\mathbf{r})v_{ext}(\mathbf{r})d\mathbf{r} + F_{HK}(\rho) \\ &= \int \rho(\mathbf{r})v_{ext}(\mathbf{r})d\mathbf{r} + T(\rho) + V_{ee}(\rho), \end{aligned} \quad (2.63)$$

where the Hohenberg-Kohn functional $F_{HK}(\rho)$ contains the kinetic energy $T(\rho)$ and the electron-electron interaction energy $V_{ee}(\rho)$

$$V_{ee}(\rho) = J(\rho) + E_{xc}(\rho). \quad (2.64)$$

Beside the Coulomb energy $J(\rho)$ it also contains a non-classical term, the so called exchange-correlation (XC) energy $E_{xc}(\rho)$. In order to obtain treatable expressions for the different terms, Hohenberg and Kohn had to resort to the Thomas-Fermi model, where they also introduced Dirac's exchange term for the uniform electron gas. These approximate expressions of course exhibit an error which becomes most apparent in the kinetic energy. In the work of Kohn and Sham [66] the DFT method has been brought to a next level by introducing self-consistent field equations for a non-interacting reference system.

2.6.3.2 Kohn-Sham-DFT

In their work of 1965 Kohn and Sham proposed an indirect approach to the calculation of the kinetic energy [66]. They suggested a non-interacting reference system described by orbitals $\{\phi_i\}$ which could in principle yield the exact ground state density of the interacting system

$$\hat{H}_{ref} = -\frac{1}{2} \sum_i^N \nabla_i^2 + \sum_i^N v_{eff}(\mathbf{r}_i). \quad (2.65)$$

This enables a simple calculation of $T(\rho)$ to a good accuracy in a similar way as the HF method

$$T_{ref}(\rho) = -\frac{1}{2} \sum_i^N \langle \phi_i | \nabla^2 | \phi_i \rangle, \quad (2.66)$$

where the corresponding density is

$$\rho(\mathbf{r}) = \sum_i^N |\phi_i(\mathbf{r})|^2, \quad (2.67)$$

as has been shown in the previous section 2.5. It has to be noted that $T_{ref}(\rho)$ is not the exact kinetic energy functional but only small corrections are left. So one obtains the Kohn-Sham equation which are of the same form as the HF-SCF equations in eq. 2.11

$$\left(\frac{1}{2} \nabla^2 + v_{eff}(\mathbf{r}) \right) \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}), \quad (2.68)$$

where the effective potential is

$$\begin{aligned} v_{eff}(\mathbf{r}) &= v_{ext}(\mathbf{r}) + \frac{\partial J(\rho)}{\partial \rho(\mathbf{r})} + \frac{\partial E_{xc}(\rho)}{\partial \rho(\mathbf{r})} \\ &= v_{ext}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r}). \end{aligned} \quad (2.69)$$

Since the XC energy functional in the KS scheme is defined as

$$E_{xc}(\rho) = T(\rho) - T_{ref}(\rho) + V_{ee}(\rho) - J(\rho), \quad (2.70)$$

the solution of the SCF equation in eq. 2.68 yields the exact density of the interacting system. Note that the XC energy functional contains – besides the Coulombic and exchange correlation – a correction to the approximated kinetic energy functional $T_{ref}(\rho)$.

Unfortunately the exact form of the XC energy functional $E_{xc}(\rho)$ is not known, but all established ab initio packages contain a variety of different approximations^{e)} which will not be presented in detail here. Compared to the MO-based correlation methods the KS-DFT method provides a favorable scaling behavior ($\mathcal{O}(M)$, see upcoming section) combined with quite accurate results. Furthermore, since the correlation effects are transferred to the XC functional, this method is not as strongly influenced by the expansion in a finite basis set as the Post-HF methods are. As will also be shown in the QMC section treating correlated wave functions, the convergence of the MO-based calculations with respect to cusp conditions is slow with the basis set size [67]. On the other hand, there exists no hierarchy of DFT methods as presented for the configuration space methods in Fig. 2.4 and with that no systematic way to approach the exact solution of the Schrödinger equation.

2.6.3.3 Time-Dependent Systems: Runge-Gross Theorem

Hohenberg and Kohn provided a fundamental existence theorem for the ground state of an electronic system based on the minimum energy principle which is not provided for time-dependent systems. Nearly twenty years later Runge and Gross [68] extended this theory to arbitrary time-dependent systems by establishing an analogous $\rho(\mathbf{r}, t) \rightarrow v_{ext}(\mathbf{r}, t)$ mapping, which leads in principle to a complete, exact density functional theory^{f)}.

^{e)}To this day the available functionals can be classified as LSDA, GGA, meta-GGA, or hybrid-type functionals. See Ref. [69] for a brief introduction to the different functional types.

^{f)}There exists of course also a time-dependent Kohn-Sham version.

With this at hand the calculation of time-dependent properties became possible within the time-dependent DFT (TDDFT) scheme which has been proven to yield reasonably accurate results for many properties [70–73]. It has to be noted that there exists a time-paradox [74] in the TDDFT equations, which will be briefly discussed in section 3.7.1.1.

2.6.3.4 Formation of Exchange-Correlation Matrices

In this section, the $\mathcal{O}(M)$ formation of the XC potential matrix \mathbf{V}_{xc} in a given basis [75] is described. It has to be noted that hybrid XC functionals [76] also contain a certain amount of exact exchange \mathbf{K} , that can be formed in $\mathcal{O}(M)$ fashion within e.g. the LinK scheme [8].

The XC energy E_{xc} is in general a functional of the density ρ and – for GGA and meta-GGA – its gradient $\nabla\rho$ and kinetic energy density τ , respectively:

$$E_{xc} = \int f_{xc}[\rho_\alpha(\mathbf{r}), \rho_\beta(\mathbf{r}), \nabla\rho_\alpha(\mathbf{r}), \nabla\rho_\beta(\mathbf{r}), \tau_\alpha(\mathbf{r}), \tau_\beta(\mathbf{r})] d\mathbf{r}. \quad (2.71)$$

The potential v_{xc} arising from exchange-correlation interactions between electrons is defined by the derivative of the XC energy functional E_{xc} with respect to the one-particle density $\rho(\mathbf{r})$

$$v_{xc}(\mathbf{r}) = \frac{f_{xc}(\mathbf{r})}{\partial\rho(\mathbf{r})}. \quad (2.72)$$

The discrete representation in a given basis results from integration over \mathbf{r}

$$\langle\chi_\mu|\hat{v}_{xc}|\chi_\nu\rangle = \int \frac{f_{xc}(\mathbf{r})}{\partial\rho(\mathbf{r})} \chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r}) d\mathbf{r}. \quad (2.73)$$

Since it is generally not possible to determine E_{xc} and \mathbf{V}_{xc} by analytic integration, a numerical quadrature has to be used, where Eq. 2.71 is rewritten as

$$E_{xc} = \sum_A^{N_A} \sum_i^{N_{grid}^A} p_A w_i f_{xc}(\mathbf{r}_i). \quad (2.74)$$

N_{grid}^A denotes the number of grid points, w_i is the weight to the given grid point \mathbf{r}_i of atom A and p_A is the nuclear partition function that enables a split of the molecular grid into single atomic integral contributions. In a first step the atomic grids are constructed usually by a combination of radial and angular grids [77]. After determination of the partition factors p_A with e.g. the popular method proposed by Becke [78] the single atomic grids are merged to yield the molecular grid in $\mathcal{O}(M)$ fashion.

For each atomic grid the integral contribution is now calculated with a scaling behavior independent of system size. After determining the constant number of basis functions χ_μ significant for the current sub-grid as well as the corresponding basis function pairs $\chi_\mu\chi_\nu$, the representation of the one-particle density within the partial grid is formed by

$$\rho(\mathbf{r}_i) = \sum_{\mu\nu} P_{\mu\nu} \chi_\mu(\mathbf{r}_i) \chi_\nu(\mathbf{r}_i), \quad (2.75)$$

with analogous equations for $\nabla\rho(\mathbf{r}_i)$ and $\tau(\mathbf{r}_i)$. At this point it is noticeable that the localization (delocalization) of the electrons resulting in a sparse (dense) discrete density \mathbf{P} does not affect the scaling behavior of the algorithm, i.e. the strict $\mathcal{O}(M)$ scaling holds even for metallic systems as long as local basis functions are used.

The evaluation of the XC functional and its derivatives at each point of the sub-grid is followed by the summation of the zeroth order values to yield the XC energy E_{xc} . To form the matrix representation of the corresponding XC potential \mathbf{V}_{xc} in the given basis, the different first order derivatives have to be contracted with the corresponding basis function values

$$\langle \chi_\mu | \hat{v}_{xc} | \chi_\nu \rangle = \sum_A^{N_A} \sum_i^{N_{grid}^A} p_A w_i \frac{\partial f_{xc}(\mathbf{r}_i)}{\partial \rho} \chi_\mu(\mathbf{r}_i) \chi_\nu(\mathbf{r}_i). \quad (2.76)$$

In a similar fashion higher-order derivatives of the XC potential necessary for response properties can be implemented, also providing a strict $\mathcal{O}(M)$ scaling behavior.

Chapter 3

Linear Scaling Density Matrix-based Methods for Static and Dynamic Properties

The theoretical determination of molecular properties for a given state (static) as well as those characterized by electronic transitions between different states (dynamic) provides a direct link between theory and experiment. The calculation of many properties is nowadays routine within the coupled-perturbed SCF scheme (for an overview see e.g. Ref. [79]), but both at the HF and KS-DFT level the size of treatable molecular systems has been constrained to the hundred atoms region due to the strong increase of the computational effort, which is often stronger than for the calculation of energies.

After an introduction to static and dynamic perturbations as well as to the corresponding responses of the molecular systems in form of the perturbed densities, a brief review of the standard MO-based schemes to calculate static (MO-CPSCF) and dynamic (MO-TDSCF) properties is given. In the main part of this chapter our new coupled-perturbed SCF equations within a local atomic orbital density matrix-based scheme for the calculation of static (D-CPSCF) and dynamic (D-TDSCF) properties are derived. The performance of their actual implementation for the calculation of NMR shielding tensors [80] and frequency-dependent polarizabilities and first order hyperpolarizabilities in the Q-Chem package [40] is presented with first exemplary applications.

3.1 Perturbative Expansion of the Energy

In the case of a weak perturbation x we can expand the energy in a Taylor series around the unperturbed case $x = 0$

$$E = E(x=0) + \left. \frac{\partial E}{\partial x} \right|_{x=0} x + \frac{1}{2} \left. \frac{\partial^2 E}{\partial x^2} \right|_{x=0} x^2 + \frac{1}{6} \left. \frac{\partial^3 E}{\partial x^3} \right|_{x=0} x^3 + \dots \quad (3.1)$$

Since molecular properties determine the response of the system to a perturbation x , we can identify them with the respective energy derivatives based on the underlying physics. For an external electric field \mathcal{E} as a perturbation we have, for example:

Dipole moment	$-\frac{dE}{d\mathcal{E}_i}$,
Polarizability	$-\frac{d^2 E}{d\mathcal{E}_i d\mathcal{E}_j}$,
1st Hyperpolarizability	$-\frac{d^3 E}{d\mathcal{E}_i d\mathcal{E}_j d\mathcal{E}_k}$.

As can be seen from the explicit expressions of the derivatives in the following section, the second and higher-order derivatives require the determination of the response of the electron distribution with respect to the external perturbation. These changes in the electron distribution destroy self-consistency in eq. 2.16, thus it has to be reinstalled in each order of the expansion in eq. 3.1 by the use of perturbation theory. From a mathematical point of view this method determines the coefficients of the Taylor expansion for a given Hamiltonian and therewith the n th order derivatives of the energy.

In contrast to the traditional approach of perturbation theory, in which the perturbed wavefunction is represented as a linear combination of the eigenfunctions of the unperturbed system, we determine the perturbed density by a variational perturbation method. In the following sections we will derive the explicit equations to determine the energy derivatives. After the derivation of an analytical expression for the first order energy derivatives, we will focus on the calculation of the n th order perturbed wavefunction or – equivalently – density which is required to determine the $(2n + 1)$ th order properties (Wigner’s rule). Their calculation in a CPSCF algorithm, which is the most time consuming step, will be treated in detail for selected properties in sections 3.6 and 3.7.

3.2 First Order Derivatives

The Hellmann-Feynman theorem states, that the first order derivative of the expectation value equals the expectation value of the derivative of the Hamiltonian

$$\frac{dE}{dx} = \langle \Psi | \frac{d\hat{H}}{dx} | \Psi \rangle. \quad (3.2)$$

However, this equation only holds – apart from some exceptions – for wavefunctions that are expanded in a complete basis, so that we cannot use this approach and have to calculate the energy derivative explicitly. We start with the derivative of eq. 2.19

$$\frac{dE}{dx} = \text{Tr} \left[\mathbf{P}^x \mathbf{h} + \mathbf{P} \mathbf{h}^x + \frac{1}{2} [\mathbf{P}^x \mathbf{G}(\mathbf{P}) + \mathbf{P} \mathbf{G}^x(\mathbf{P}) + \mathbf{P} \mathbf{G}(\mathbf{P}^x)] \right], \quad (3.3)$$

where \mathbf{h} is the matrix representation of the core-Hamiltonian and \mathbf{G} contains the Coulomb and exchange matrices. Rearranging this equation slightly and considering the identity $\text{Tr} [\mathbf{P} \mathbf{G}(\mathbf{P}^x)] = \text{Tr} [\mathbf{P}^x \mathbf{G}(\mathbf{P})]$ we get

$$\begin{aligned} \frac{dE}{dx} &= \text{Tr} \left[\mathbf{P} \mathbf{h}^x + \frac{1}{2} \mathbf{P} \mathbf{G}^x(\mathbf{P}) \right] + \text{Tr} \left[\mathbf{P}^x \mathbf{h} + \mathbf{P}^x \mathbf{G}(\mathbf{P}) \right] \\ &= \text{Tr} \left[\mathbf{P} \mathbf{h}^x + \frac{1}{2} \mathbf{P} \mathbf{G}^x(\mathbf{P}) \right] + \text{Tr} \left[\mathbf{P}^x \mathbf{F} \right]. \end{aligned} \quad (3.4)$$

The terms of the first trace are computed with standard algorithms in a linear scaling manner, but the second trace contains the perturbed density. In order to obtain a result independent of \mathbf{P}^x [81–83], we have to anticipate the properties of the perturbed density described in section 3.3. The resolution of \mathbf{P}^x into a sum of projections shows that the virt/virt part \mathbf{P}^x_{vv} vanishes and $\mathbf{P}^x_{\text{oo}} = -\mathbf{P} \mathbf{S}^x \mathbf{P}$. From the HF- or KS-equations we see that the MO representation of the matrix \mathbf{F} is in diagonal form and so has only projections onto the pure occupied and virtual subspaces. Analogously, the commutator in eq. 2.46 and the idempotency relation in eq. 2.40 shows that the projections of the Fock matrix on the occ/virt and virt/occ subspaces vanish

$$\begin{aligned} \mathbf{F}_{\text{ov}} &= \mathbf{S} \mathbf{P} \mathbf{F} (\mathbf{1} - \mathbf{P} \mathbf{S}) &= \mathbf{S} \mathbf{P} \mathbf{F} - \mathbf{S} \mathbf{P} \mathbf{F} \mathbf{P} \mathbf{S} \\ &= \mathbf{S} \mathbf{P} \mathbf{S} \mathbf{P} \mathbf{F} - \mathbf{S} \mathbf{P} \mathbf{F} \mathbf{P} \mathbf{S} = \mathbf{S} \mathbf{P} \mathbf{F} \mathbf{P} \mathbf{S} - \mathbf{S} \mathbf{P} \mathbf{F} \mathbf{P} \mathbf{S} = \mathbf{0}, \end{aligned} \quad (3.5)$$

$$\begin{aligned} \mathbf{F}_{\text{vo}} &= (\mathbf{1} - \mathbf{S} \mathbf{P}) \mathbf{F} \mathbf{P} \mathbf{S} &= \mathbf{F} \mathbf{P} \mathbf{S} - \mathbf{S} \mathbf{P} \mathbf{F} \mathbf{P} \mathbf{S} \\ &= \mathbf{F} \mathbf{P} \mathbf{S} \mathbf{P} \mathbf{S} - \mathbf{S} \mathbf{P} \mathbf{F} \mathbf{P} \mathbf{S} = \mathbf{S} \mathbf{P} \mathbf{F} \mathbf{P} \mathbf{S} - \mathbf{S} \mathbf{P} \mathbf{F} \mathbf{P} \mathbf{S} = \mathbf{0}. \end{aligned} \quad (3.6)$$

Since the trace is invariant with respect to cyclic permutations of its arguments

$$\text{Tr} [\mathbf{P}^x_{\text{ov}} \mathbf{F}] = \text{Tr} [\mathbf{PSP}^x(\mathbf{1} - \mathbf{SP})\mathbf{F}] = \text{Tr} [\mathbf{P}^x(\mathbf{1} - \mathbf{SP})\mathbf{FPS}] = \mathbf{0}, \quad (3.7)$$

$$\text{Tr} [\mathbf{P}^x_{\text{vo}} \mathbf{F}] = \text{Tr} [(\mathbf{1} - \mathbf{PS})\mathbf{P}^x\mathbf{SPF}] = \text{Tr} [\mathbf{P}^x\mathbf{SPF}(\mathbf{1} - \mathbf{PS})] = \mathbf{0}, \quad (3.8)$$

one obtains

$$\begin{aligned} \frac{dE}{dx} &= \text{Tr} \left[\mathbf{Ph}^x + \frac{1}{2} \mathbf{PG}^x(\mathbf{P}) \right] - \text{Tr} [\mathbf{PS}^x \mathbf{PF}] \\ &= \text{Tr} \left[\mathbf{Ph}^x + \frac{1}{2} \mathbf{PG}^x(\mathbf{P}) \right] - \text{Tr} [\mathbf{WS}^x], \end{aligned} \quad (3.9)$$

with $\mathbf{W} = \mathbf{PFP}$ as the energy-weighted density matrix. An $\mathcal{O}(M)$ algorithm of first order properties can be easily implemented by an extension of the Link/CFMM schemes [1, 8] to the perturbed two-electron integrals $\mathbf{G}^x(\mathbf{P})$ [5, 9].

The second derivative of the energy follows from eq. 3.9 as

$$\begin{aligned} \frac{d^2 E}{dx dy} &= \text{Tr} \left[\mathbf{Ph}^{xy} + \frac{1}{2} \mathbf{PG}^{xy}(\mathbf{P}) - \mathbf{WS}^{xy} + \mathbf{PF}^y \mathbf{PS}^x \right] \\ &+ \text{Tr} \left[\mathbf{P}^y \mathbf{h}^x + \mathbf{P}^y \mathbf{G}^x(\mathbf{P}) - \mathbf{P}^y \mathbf{FPS}^x + \mathbf{PFP}^y \mathbf{S}^x \right]. \end{aligned} \quad (3.10)$$

The determination of the derivative of the density in the second trace cannot be avoided anymore and has to be obtained as solution from the CPSCF equations.

3.3 Properties of the First-Order Perturbed Density

The unique resolution described in eq. 2.50 gives for the perturbed density matrix

$$\mathbf{P}^x = \mathbf{P}^x_{\text{oo}} + \mathbf{P}^x_{\text{ov}} + \mathbf{P}^x_{\text{vo}} + \mathbf{P}^x_{\text{vv}}. \quad (3.11)$$

From the derivative of the idempotency condition eq. 2.40

$$\mathbf{P}^x = \mathbf{P}^x \mathbf{SP} + \mathbf{PS}^x \mathbf{P} + \mathbf{PSP}^x, \quad (3.12)$$

we can rewrite the terms in the resolution of \mathbf{P}^x by projecting eq. 3.12 onto the subspaces:

$$\mathbf{P}^x_{\text{oo}} = -\mathbf{PS}^x \mathbf{P}, \quad (3.13)$$

$$\mathbf{P}^x_{\text{ov}} = \mathbf{PSP}^x - \mathbf{PSP}^x \mathbf{SP}, \quad (3.14)$$

$$\mathbf{P}^x_{\text{vo}} = \mathbf{P}^x \mathbf{SP} - \mathbf{PSP}^x \mathbf{SP}, \quad (3.15)$$

$$\mathbf{P}^x_{\text{vv}} = \mathbf{0}. \quad (3.16)$$

The hermiticity in eq. 2.47 has also to be fulfilled for the derivatives ($\mathbf{P}^x = \mathbf{P}^{x\dagger}$), so we obtain

$$\mathbf{P}_{\text{ov}}^x = \mathbf{P}_{\text{vo}}^{x\dagger}. \quad (3.17)$$

Since we can directly calculate \mathbf{S}^x we only have to determine the occ/virt part \mathbf{P}_{ov}^x . Note that eqs. 3.12-3.17 only hold for the converged perturbed densities, during the iterations all constraints have to be imposed by the properties of the projection operators which is discussed in the corresponding sections.

3.4 Frequency-Dependent Perturbations and Properties of the First and Second Order Transition Densities

For the description of fluctuations of arbitrary operators in time caused by the presence of a time-dependent oscillatory perturbation $\hat{H}^{(S)}(t)$ we start from the time-dependent Schrödinger equation (eq. 2.1) in atomic units:

$$\hat{H} |\Psi\rangle = i \frac{\partial}{\partial t} |\Psi\rangle, \quad (3.18)$$

where the Hamilton operator is the sum of the time-independent Hamiltonian $\hat{H}^{(0)}$ of the unperturbed molecular system and a time-dependent perturbation:

$$\hat{H} = \hat{H}^{(0)} + \hat{H}^{(S)}(t). \quad (3.19)$$

In the case of molecular polarizabilities the perturbation $\hat{H}^{(S)}$ of an external monochromatic oscillatory electric field is represented by the interaction of the molecular system with a single Fourier component of the quantized photon field within the dipole approximation

$$\hat{H}^{(S)} = \frac{1}{2} \hat{\mu} \mathcal{E} (e^{-i\omega t} + e^{+i\omega t}), \quad (3.20)$$

$$\hat{\mu} = -z_{el} \sum_{i=1}^N \hat{r}_i, \quad (3.21)$$

where \mathcal{E} is the electric field vector, ω the corresponding frequency, z_{el} the electronic charge and $\hat{\mu}$ the dipole moment operator.

Table 3.1: Polarizabilities and first order hyperpolarizabilities. κ is the prefactor of the RHS of the quadratic response equations (eq. 3.163).

	Property		κ
Static polarizability:	$\alpha_{xy}(0)$	$= -\text{Tr} [\mathbf{h}^{\mu x} \mathbf{P}^y (\omega = 0)]$	0
Dynamic polarizability:	$\alpha_{xy}(\mp\omega; \pm\omega)$	$= -\text{Tr} [\mathbf{h}^{\mu x} \mathbf{P}^y (\pm\omega)]$	∓ 1
First static hyperpolarizability:	$\beta_{xyz}(0; 0, 0)$	$= -\text{Tr} [\mathbf{h}^{\mu x} \mathbf{P}^{yz} (\omega = 0, \omega = 0)]$	0
2nd harmonic generation:	$\beta_{xyz}(\mp 2\omega; \pm\omega, \pm\omega)$	$= -\text{Tr} [\mathbf{h}^{\mu x} \mathbf{P}^{yz} (\pm\omega \pm \omega)]$	∓ 2
Electro-optical Pockel's effect:	$\beta_{xyz}(\mp\omega; 0, \pm\omega)$	$= -\text{Tr} [\mathbf{h}^{\mu x} \mathbf{P}^{yz} (0, \pm\omega)]$	∓ 1
Optical rectification:	$\beta_{xyz}(0; \pm\omega, \mp\omega)$	$= -\text{Tr} [\mathbf{h}^{\mu x} \mathbf{P}^{yz} (\pm\omega, \mp\omega)]$	0

First of all we want to identify the optical properties from the expansion of the molecular polarizability \mathcal{P} as a power series in the strength of the external electric field around the field-free state. So we have to choose $\partial E/\partial \mathcal{E}$ as the starting point of the expansion:

$$\begin{aligned}
 \mathcal{P}_i &= \left. \frac{\partial E}{\partial \mathcal{E}_i} + \frac{\partial^2 E}{\partial \mathcal{E}_i \partial \mathcal{E}_j} \right|_{\mathcal{E}_j=0} \mathcal{E}_j + \frac{1}{2} \left. \frac{\partial^3 E}{\partial \mathcal{E}_i \partial \mathcal{E}_j \partial \mathcal{E}_k} \right|_{\mathcal{E}_j=\mathcal{E}_k=0} \mathcal{E}_j \mathcal{E}_k + \dots \\
 &= -\mu_i - \alpha_{ij} \mathcal{E}_j - \frac{1}{2} \beta_{ijk} \mathcal{E}_j \mathcal{E}_k + \dots, \tag{3.22}
 \end{aligned}$$

where $i, j, k = x, y, z$ and summation over same indices is assumed. Considering the different terms in the perturbation eq. 3.20 the properties listed in Tab. 3.1 can be identified.

The structure of the 1st order transition density $\mathbf{P}^x(\pm\omega)$ can be investigated starting with the conditions given in eq. 2.42 and eq. 2.47. From the derivative of eq. 2.47 it follows that

$$[\mathbf{P}^x(+\omega) e^{+i\omega t} + \mathbf{P}^x(-\omega) e^{-i\omega t}]^\dagger = [(\mathbf{P}^x(+\omega))^\dagger e^{-i\omega t} + (\mathbf{P}^x(-\omega))^\dagger e^{+i\omega t}], \tag{3.23}$$

$$\implies \mathbf{P}^x(+\omega) = (\mathbf{P}^x(-\omega))^\dagger,$$

$$\mathbf{P}^x(+\omega) \neq (\mathbf{P}^x(+\omega))^\dagger \quad \text{for } \omega \neq 0, \tag{3.24}$$

i.e. the 1st order density is unsymmetric for the dynamic case and in the limit ($\omega \rightarrow 0$) we have a hermetic \mathbf{P}^x as expected for the static case. Nevertheless, we only have to solve for one time factor (excitation or de-excitation) and get the other by the transposition. For

simplicity we ignore the time factor of the density expansion since the following relations will hold for all the terms. From the derivative of the idempotency condition in eq. 3.12 we can get in contact with the terms in the resolution (eq. 2.50) of \mathbf{P}^x :

$$\mathbf{P}_{oo}^x = \mathbf{0}, \quad (3.25)$$

$$\mathbf{P}^x \mathbf{S} \mathbf{P} = \mathbf{P}_{vo}^x, \quad (3.26)$$

$$\mathbf{P} \mathbf{S} \mathbf{P}^x = \mathbf{P}_{ov}^x, \quad (3.27)$$

$$\mathbf{P}_{vv}^x = \mathbf{0}, \quad (3.28)$$

where the first equation results from the fact that the basis functions do not depend on the perturbation ($\mathbf{S}^x = 0$).

For the investigation of the 2nd order density it is also useful to consider the spectral representation of \mathbf{P}^x . From the expansion of a general wavefunction in the stationary eigenfunctions (see e.g. Ref. [32]) we obtain in an analogous way, i.e. by expansion in the basis of one-electron molecular spin-orbital products, to eq. 2.30 (skipping the time factor)

$$P_{\mu\nu}^x = \sum_{\substack{i \in occ \\ a \in virt}} x_a^i C_{\mu a} C_{\nu i}^* + \sum_{\substack{i \in occ \\ a \in virt}} y_i^a C_{\mu i} C_{\nu a}^*. \quad (3.29)$$

The coefficients x_a^i and y_i^a are the virt/occ and occ/virt parts of the matrix representation of the transition density in the MO basis. They can be interpreted as transitions for de-excitation (x_a^i : $\phi_a \rightarrow \phi_i$) and excitation (y_i^a : $\phi_i \rightarrow \phi_a$).

The second derivative of the idempotency relation for all cases of \mathbf{P}^{xy} , i.e. all frequency combinations of x and y in β_{xyz} as shown in Tab. 3.1, is

$$\mathbf{P}^{xy} = \mathbf{P}^{xy} \mathbf{S} \mathbf{P} + \mathbf{P}^x \mathbf{S} \mathbf{P}^y + \mathbf{P}^y \mathbf{S} \mathbf{P}^x + \mathbf{P} \mathbf{S} \mathbf{P}^{xy}. \quad (3.30)$$

First, the projections of the density onto the occupied subspace are investigated by multiplying eq. 3.30 with \mathbf{P}_{occ} :

$$\mathbf{P} \mathbf{S} \mathbf{P}^{xy} = \mathbf{P} \mathbf{S} \mathbf{P}^{xy} \mathbf{S} \mathbf{P} + \mathbf{P} \mathbf{S} \mathbf{P}^x \mathbf{S} \mathbf{P}^y + \mathbf{P} \mathbf{S} \mathbf{P}^y \mathbf{S} \mathbf{P}^x + \mathbf{P} \mathbf{S} \mathbf{P} \mathbf{S} \mathbf{P}^{xy}, \quad (3.31)$$

$$\mathbf{P}^{xy} \mathbf{S} \mathbf{P} = \mathbf{P}^{xy} \mathbf{S} \mathbf{P} \mathbf{S} \mathbf{P} + \mathbf{P}^x \mathbf{S} \mathbf{P}^y \mathbf{S} \mathbf{P} + \mathbf{P}^y \mathbf{S} \mathbf{P}^x \mathbf{S} \mathbf{P} + \mathbf{P} \mathbf{S} \mathbf{P}^{xy} \mathbf{S} \mathbf{P}, \quad (3.32)$$

$$\begin{aligned} \mathbf{P} \mathbf{S} \mathbf{P}^{xy} \mathbf{S} \mathbf{P} &= \mathbf{P} \mathbf{S} \mathbf{P}^{xy} \mathbf{S} \mathbf{P} \mathbf{S} \mathbf{P} + \mathbf{P} \mathbf{S} \mathbf{P}^x \mathbf{S} \mathbf{P}^y \mathbf{S} \mathbf{P} + \mathbf{P} \mathbf{S} \mathbf{P}^y \mathbf{S} \mathbf{P}^x \mathbf{S} \mathbf{P} \\ &\quad + \mathbf{P} \mathbf{S} \mathbf{P} \mathbf{S} \mathbf{P}^{xy} \mathbf{S} \mathbf{P}. \end{aligned} \quad (3.33)$$

Using the idempotency relation eq. 2.42, the three equations above lead to the following relations for the occ/occ part of \mathbf{P}^{xy} :

$$\begin{aligned}\mathbf{PSP}^{xy}\mathbf{SP} &= -\mathbf{PSP}^x\mathbf{SP}^y - \mathbf{PSP}^y\mathbf{SP}^x \\ &= -\mathbf{P}^x\mathbf{SP}^y\mathbf{SP} - \mathbf{P}^y\mathbf{SP}^x\mathbf{SP} \\ &= -\mathbf{PSP}^x\mathbf{SP}^y\mathbf{SP} - \mathbf{PSP}^y\mathbf{SP}^x\mathbf{SP}.\end{aligned}\tag{3.34}$$

From eqs. 3.31-3.34 we get for the projections of \mathbf{P}^{xy} onto the subspaces

$$\begin{aligned}\mathbf{P}_{oo}^{xy} &= \mathbf{PSP}^x\mathbf{SP}^y\mathbf{SP} + \mathbf{PSP}^y\mathbf{SP}^x\mathbf{SP} + 2\mathbf{PSP}^{xy}\mathbf{SP} \\ &= -\mathbf{PSP}^x\mathbf{SP}^y\mathbf{SP} - \mathbf{PSP}^y\mathbf{SP}^x\mathbf{SP},\end{aligned}\tag{3.35}$$

$$\mathbf{P}_{ov}^{xy} = \mathbf{PSP}^{xy} - \mathbf{PSP}^{xy}\mathbf{SP},\tag{3.36}$$

$$\mathbf{P}_{vo}^{xy} = \mathbf{P}^{xy}\mathbf{SP} - \mathbf{PSP}^{xy}\mathbf{SP},\tag{3.37}$$

$$\begin{aligned}\mathbf{P}_{vv}^{xy} &= \mathbf{P}^x\mathbf{SP}^y + \mathbf{P}^y\mathbf{SP}^x - \mathbf{PSP}^x\mathbf{SP}^y - \mathbf{PSP}^y\mathbf{SP}^x \\ &\quad - \mathbf{P}^x\mathbf{SP}^y\mathbf{SP} - \mathbf{P}^y\mathbf{SP}^x\mathbf{SP} + \mathbf{PSP}^x\mathbf{SP}^y\mathbf{SP} + \mathbf{PSP}^y\mathbf{SP}^x\mathbf{SP} \\ &= \mathbf{P}^x\mathbf{SP}^y + \mathbf{P}^y\mathbf{SP}^x - \mathbf{PSP}^{xy}\mathbf{SP} \\ &= \mathbf{P}^x\mathbf{SPSP}^y + \mathbf{P}^y\mathbf{SPSP}^x,\end{aligned}\tag{3.38}$$

$$\implies \mathbf{P}_{vv}^{xy} + \mathbf{P}_{oo}^{xy} = \mathbf{P}^x\mathbf{SP}^y + \mathbf{P}^y\mathbf{SP}^x.\tag{3.39}$$

Again we only have to solve for \mathbf{P}_{ov}^{xy} and \mathbf{P}_{vo}^{xy} since the occ/occ and virt/virt parts are completely determined by the 1st order transition densities (eq. 3.39). The resolution of the term $\mathbf{P}^x\mathbf{SP}^y$ gives

$$(\mathbf{P}^x\mathbf{SP}^y)_{oo} = \mathbf{PSP}^x\mathbf{SP}^y\mathbf{SP},\tag{3.40}$$

$$(\mathbf{P}^x\mathbf{SP}^y)_{vv} = \mathbf{P}^x\mathbf{SPSP}^y,\tag{3.41}$$

$$(\mathbf{P}^x\mathbf{SP}^y)_{ov} = (\mathbf{P}^x\mathbf{SP}^y)_{vo} = \mathbf{0},\tag{3.42}$$

which is evident from the spectral representation of \mathbf{P}^x in eq. 3.29 considering the orthonormality of the MOs in eq. 2.41.

3.5 MO-based CPSCF/TDSCF Equations

In this section the molecular orbital-based coupled-perturbed SCF (MO-CPSCF) equations are briefly discussed, as they have been implemented within this work in order to compare

their results with those of the new methods. Starting with the more general time-dependent equations, the static case is derived in the frequency limit $\omega \rightarrow 0$.

3.5.1 Dynamic Properties: MO-TDSCF

In this section the derivation of the TDSCF equations will closely follow the work of Sekino and Bartlett [84]. Starting from Frenkel's variational principle [47, 85] the time-dependent Roothaan-Hall equations are

$$\mathbf{FC} - \mathbf{SC}\epsilon = i\frac{\partial}{\partial t}\mathbf{SC}. \quad (3.43)$$

Invoking the perturbative expansion for the Fock matrix we obtain

$$\begin{aligned} \mathbf{F} &= \mathbf{F}^{(0)} + \mathcal{E} \left(e^{-i\omega t}\mathbf{F}^{(1)}(-\omega) + e^{+i\omega t}\mathbf{F}^{(1)}(+\omega) + \mathbf{F}^{(1)}(\omega = 0) \right) \\ &+ \mathcal{E}^2 \left(e^{-2i\omega t}\mathbf{F}^{(2)}(-\omega, -\omega) + e^{+2i\omega t}\mathbf{F}^{(2)}(+\omega, +\omega) + 2\mathbf{F}^{(2)}(+\omega, -\omega) + \mathbf{F}^{(2)}(0, 0) \right) \\ &+ \dots \end{aligned} \quad (3.44)$$

and similar expressions for \mathbf{C} , \mathbf{P} , ϵ and \mathbf{S} . Note that the basis functions do not depend on the external electric field, so we will only have to resort to an expansion of the metric in the case of static perturbations like e.g. nuclear displacements. Furthermore, the derivatives of the integrals also vanish and only terms of Coulomb and exchange type contractions with transition densities of corresponding order occur. The derivatives of the Fock matrix are

$$\mathbf{F}^x(\pm\omega) = \mathbf{h}^{\mu_x} + \mathbf{G}(\mathbf{P}^x(\pm\omega)), \quad (3.45)$$

$$\mathbf{F}^{xy}(\pm\omega, \pm\omega) = \mathbf{G}(\mathbf{P}^{xy}(\pm\omega, \pm\omega)), \quad (3.46)$$

while the perturbed molecular orbital coefficients are expanded in the basis of the unperturbed solution

$$\mathbf{C}^x(\pm\omega) = \mathbf{C}^{(0)}\mathbf{U}^x(\pm\omega), \quad \mathbf{C}^{xy}(\pm\omega, \pm\omega) = \mathbf{C}^{(0)}\mathbf{U}^{xy}(\pm\omega, \pm\omega), \quad (3.47)$$

with \mathbf{U} as the transition coefficients matrix. Similar second order expressions exist for the frequency combinations $(\pm\omega, \mp\omega)$ and $(\pm\omega, \omega = 0)$. In order to explain the general principle we will focus on the calculation of second harmonic generation (SHG) in the following text.

Here the CPSCF equations are solved in the MO basis, so the transition densities have to be formed explicitly from the MO coefficients

$$\mathbf{P}^x(\pm\omega) = \mathbf{C}^x(\pm\omega)\mathbf{P}^{MO}\mathbf{C}^{(0)\dagger} + \mathbf{C}^{(0)}\mathbf{P}^{MO}\mathbf{C}^{x\dagger}(\mp\omega), \quad (3.48)$$

$$\begin{aligned} \mathbf{P}^{xy}(\pm\omega, \pm\omega) &= \mathbf{C}^{xy}(\pm\omega, \pm\omega)\mathbf{P}^{MO}\mathbf{C}^{(0)\dagger} + \mathbf{C}^x(\pm\omega)\mathbf{P}^{MO}\mathbf{C}^{y\dagger}(\mp\omega) \\ &+ \mathbf{C}^y(\pm\omega)\mathbf{P}^{MO}\mathbf{C}^{x\dagger}(\mp\omega) + \mathbf{C}^{(0)}\mathbf{P}^{MO}\mathbf{C}^{xy\dagger}(\pm\omega, \pm\omega), \end{aligned} \quad (3.49)$$

where \mathbf{P}^{MO} is the diagonal matrix containing the occupation numbers of the different molecular orbitals.

3.5.1.1 First-Order MO-TDSCF Equations

Inserting the different expansions for \mathbf{F} , \mathbf{C} and ϵ into eq. 3.43 and sorting the terms by orders yields for the first order

$$\mathbf{F}^x(\pm\omega)\mathbf{C}^{(0)} + \mathbf{F}^{(0)}\mathbf{C}^x(\pm\omega) - \mathbf{S}\mathbf{C}^x(\pm\omega)\epsilon^{(0)} - \mathbf{S}\mathbf{C}\epsilon^x(\pm\omega) = \mp\omega\mathbf{S}\mathbf{C}^x(\pm\omega), \quad (3.50)$$

where the derivative with respect to time t has been evaluated already. Multiplying eq. 3.50 from left with $\mathbf{C}^{(0)}$, inserting the resulting MO representations ($\mathbf{C}^{(0)\dagger}\mathbf{A}\mathbf{C}^{(0)} = \mathbf{A}_{MO}$) and considering the orthonormality constraint ($\mathbf{C}^{(0)\dagger}\mathbf{S}\mathbf{C}^{(0)} = \mathbf{1}$) we obtain after a slight rearrangement of the terms

$$\epsilon^x(\pm\omega) = \mathbf{F}_{MO}^x(\pm\omega) + (\epsilon^{(0)} \pm\omega)\mathbf{U}^x(\pm\omega) - \mathbf{U}^x(\pm\omega)\epsilon^{(0)}. \quad (3.51)$$

Because of the strict orthogonality between the occupied and virtual subspaces, the matrix $\epsilon^x(\pm\omega)$ is at least block-diagonal, i.e. $\epsilon_{rs}^x(\pm\omega) = 0$ for $(r \in occ, s \in virt)$ and $(r \in virt, s \in occ)$, respectively. The occ/virt and virt/occ elements of the transition coefficients matrix result as

$$\mathbf{U}_{ia}^x(\pm\omega) = \frac{\mathbf{F}_{MO,ia}^x(\pm\omega)}{\epsilon_a^{(0)} - \epsilon_i^{(0)} \mp\omega}. \quad (3.52)$$

From the first order derivative of the orthonormality constraint we can obtain a connection between $\mathbf{U}(+\omega)$ and $\mathbf{U}(-\omega)$

$$\begin{aligned} \mathbf{C}^{(0)\dagger}\mathbf{S}\mathbf{C}^x(\pm\omega) + \mathbf{C}^{x\dagger}(\mp\omega)\mathbf{S}\mathbf{C}^{(0)} &= \mathbf{C}^{(0)\dagger}\mathbf{S}\mathbf{C}^{(0)}\mathbf{U}^x(\pm\omega) + \mathbf{U}^{x\dagger}(\mp\omega)\mathbf{C}^{(0)\dagger}\mathbf{S}\mathbf{C}^{(0)} \\ &= \mathbf{U}^x(\pm\omega) + \mathbf{U}^{x\dagger}(\mp\omega) \\ &= \mathbf{0}, \\ \implies \mathbf{U}^x(\pm\omega) &= -\mathbf{U}^{x\dagger}(\mp\omega). \end{aligned} \quad (3.53)$$

Thus we only have to solve the linear equation system (LEQS)

$$\begin{aligned} \mathbf{AU} &= \mathbf{b}, \\ \implies (\epsilon_a^{(0)} - \epsilon_i^{(0)} \mp \omega) \mathbf{U}_{ia}^x(\pm\omega) - \mathbf{G}_{MO,ia}(\mathbf{P}^x(\pm\omega)) &= \mathbf{H}_{MO}^{\mu_x} \end{aligned} \quad (3.54)$$

while $\mathbf{C}^x(\mp\omega)$, which is needed to form $\mathbf{P}^x(\pm\omega)$, is constructed as $\mathbf{C}^x(\mp\omega) = -\mathbf{C}^{(0)}\mathbf{U}^{x\dagger}(\pm\omega)$. Note that eq. 3.52 diverges if the frequency ω coincides with an orbital energy difference $\epsilon_j^{(0)} - \epsilon_i^{(0)}$ or for a degenerated system in the static case $\omega = 0$. In these cases the MOs of the single subspaces can be transformed to a non-canonical solution [86], avoiding divergences in eq. 3.52.

3.5.1.2 Second-Order MO-TDSCF Equations

To obtain the second order transition density to determine SHG processes, the second-order terms of the expansions of \mathbf{F} , \mathbf{S} and ϵ are inserted into the time-dependent Roothaan-Hall equation in eq. 3.43

$$\begin{aligned} &\mathbf{F}^{xy}(\pm\omega, \pm\omega) \mathbf{C}^{(0)} + \mathbf{F}^x(\pm\omega) \mathbf{C}^y(\pm\omega) + \mathbf{F}^y(\pm\omega) \mathbf{C}^x(\pm\omega) \\ &+ \mathbf{F}^{(0)} \mathbf{C}^{xy}(\pm\omega, \pm\omega) - \mathbf{S} \mathbf{C}^{xy}(\pm\omega, \pm\omega) \epsilon^{(0)} - \mathbf{S} \mathbf{C}^x(\pm\omega) \epsilon^y(\pm\omega) \\ &- \mathbf{S} \mathbf{C}^y(\pm\omega) \epsilon^x(\pm\omega) - \mathbf{S} \mathbf{C}^{(0)}(\pm\omega, \pm\omega) \epsilon^{xy}(\pm\omega, \pm\omega) \\ &= \mp 2\omega \mathbf{S} \mathbf{C}^{xy}(\pm\omega, \pm\omega). \end{aligned} \quad (3.55)$$

Proceeding in the same manner as in the first order case we transform eq. 3.55 to the MO basis and rearrange the terms to

$$\begin{aligned} \epsilon^{xy}(\pm\omega, \pm\omega) &= \mathbf{G}_{MO}(\mathbf{P}^{xy}(\pm\omega, \pm\omega)) + \mathbf{F}_{MO}^x(\pm\omega) \mathbf{U}^y(\pm\omega) \\ &+ \mathbf{F}_{MO}^y(\pm\omega) \mathbf{U}^x(\pm\omega) - \mathbf{U}^x(\pm\omega) \epsilon^y(\pm\omega) - \mathbf{U}^y(\pm\omega) \epsilon^x(\pm\omega) \\ &+ (\epsilon^{(0)} \pm 2\omega) \mathbf{U}^{xy}(\pm\omega, \pm\omega) - \mathbf{U}^{xy}(\pm\omega, \pm\omega) \epsilon. \end{aligned} \quad (3.56)$$

Collecting the following first order quantities as

$$\begin{aligned} \mathbf{T}^{xy}(\pm\omega, \pm\omega) &= \mathbf{F}_{MO}^x(\pm\omega) \mathbf{U}^y(\pm\omega) + \mathbf{F}_{MO}^y(\pm\omega) \mathbf{U}^x(\pm\omega) \\ &- \mathbf{U}^x(\pm\omega) \mathbf{F}_{MO}^y(\pm\omega) - \mathbf{U}^y(\pm\omega) \mathbf{F}_{MO}^x(\pm\omega), \end{aligned} \quad (3.57)$$

the occ/virt and virt/occ blocks of the second order transition coefficients can be written as

$$\mathbf{U}_{ia}^{xy}(\pm\omega, \pm\omega) = \frac{\mathbf{G}_{MO,ia}(\mathbf{P}^{xy}(\pm\omega, \pm\omega)) + \mathbf{T}^{xy}(\pm\omega, \pm\omega)}{\epsilon_a^{(0)} - \epsilon_i^{(0)} \mp 2\omega}. \quad (3.58)$$

Analogous to first order we get the relation $\mathbf{U}^{xy}(\pm\omega, \pm\omega) = -\mathbf{U}^{xy\dagger}(\mp\omega, \mp\omega)$ for the off-diagonal blocks, the pure occ/occ and virt/virt blocks of $\mathbf{U}^{xy}(\pm\omega, \pm\omega)$ can be exclusively written in first order quantities

$$\mathbf{U}^{xy}(\pm\omega, \pm\omega) = \frac{1}{2}(\mathbf{U}^x(\pm\omega)\mathbf{U}^y(\pm\omega) + \mathbf{U}^y(\pm\omega)\mathbf{U}^x(\pm\omega)). \quad (3.59)$$

If the first order results are provided, we can solve the resulting linear equation system (LEQS) in order to determine the third order properties given in Tab. 3.1.

3.5.1.3 Wigner (2n+1) Rule

As could be seen in the preceding section, the second order TDSCF scheme is partially constructed from first order results. So it is clear that the result as well as the convergence behavior of the second order TDSCF algorithm depend on the quality of those results. For third-order properties one could alternatively employ the Wigner (2n+1) rule which reduces the overall number of iterative calculations. In this brief section no detailed derivation is given, only the final expression for a component of β_{SHG} based on the first order results for ω and 2ω according to Karna and Dupuis [86]:

$$\begin{aligned} \beta_{xyx}(-2\omega; +\omega, +\omega) = & \text{Tr}_{\text{occ}} \left[\mathbf{U}^x(-2\omega)\mathbf{F}_{MO}^y(+\omega)\mathbf{U}^z(+\omega) + \mathbf{U}^z(+\omega)\mathbf{F}_{MO}^y(+\omega)\mathbf{U}^x(-2\omega) \right. \\ & + \mathbf{U}^y(+\omega)\mathbf{F}_{MO}^z(+\omega)\mathbf{U}^x(-2\omega) + \mathbf{U}^x(-2\omega)\mathbf{F}_{MO}^z(+\omega)\mathbf{U}^y(+\omega) \\ & + \mathbf{U}^z(+\omega)\mathbf{F}_{MO}^x(-2\omega)\mathbf{U}^y(+\omega) + \mathbf{U}^y(+\omega)\mathbf{F}_{MO}^x(-2\omega)\mathbf{U}^z(+\omega) \\ & - \mathbf{U}^x(-2\omega)\mathbf{U}^z(+\omega)\boldsymbol{\epsilon}^y(+\omega) - \mathbf{U}^z(+\omega)\mathbf{U}^x(-2\omega)\boldsymbol{\epsilon}^y(+\omega) \\ & - \mathbf{U}^y(+\omega)\mathbf{U}^x(-2\omega)\boldsymbol{\epsilon}^z(+\omega) - \mathbf{U}^x(-2\omega)\mathbf{U}^y(+\omega)\boldsymbol{\epsilon}^z(+\omega) \\ & \left. - \mathbf{U}^z(+\omega)\mathbf{U}^y(+\omega)\boldsymbol{\epsilon}^x(-2\omega) - \mathbf{U}^y(-2\omega)\mathbf{U}^z(+\omega)\boldsymbol{\epsilon}^x(-2\omega) \right]. \quad (3.60) \end{aligned}$$

Here $\text{Tr}_{\text{occ}}[\mathbf{A}]$ is the sum of the first N_{occ} diagonal elements where N_{occ} is the number of occupied orbitals. It can also be formulated by a trace of the argument multiplied with the occupation number matrix \mathbf{P}^{MO} (see sec. 2.5.1):

$$\text{Tr}_{\text{occ}}[\mathbf{A}] = \text{Tr}[\mathbf{P}^{MO}\mathbf{A}]. \quad (3.61)$$

This equation will be referred to when the corresponding density matrix-based equations are derived.

3.5.2 Static Second-Order Properties: First-Order CPSCF

If the basis functions do not depend on the external perturbation, we can simply use eq. 3.52 in the static limit $\omega \rightarrow 0$. But in most cases like vibrational frequencies or NMR shieldings using GIAOs we have to consider the derivatives of the metric because of the explicit dependence of the basis functions on the perturbation. Similar to eq. 3.50 we obtain in first order

$$\mathbf{F}^x \mathbf{C}^{(0)} + \mathbf{F}^{(0)} \mathbf{C}^x - \mathbf{S}^x \mathbf{C}^{(0)} \boldsymbol{\epsilon}^{(0)} - \mathbf{S}^{(0)} \mathbf{C}^x \boldsymbol{\epsilon}^{(0)} - \mathbf{S}^{(0)} \mathbf{C}^{(0)} \boldsymbol{\epsilon}^x = \mathbf{0}, \quad (3.62)$$

where the time-dependency on the RHS vanishes. The transformation into the MO basis leads to

$$\boldsymbol{\epsilon}^x = \mathbf{F}_{MO} + \boldsymbol{\epsilon}^{(0)} \mathbf{U}^x - \mathbf{S}_{MO}^x \boldsymbol{\epsilon} - \mathbf{U}^x \boldsymbol{\epsilon}^{(0)}, \quad (3.63)$$

from which the occ/virt and virt/occ parts of \mathbf{U}^x are obtained

$$\mathbf{U}_{ia}^x = \frac{\mathbf{F}_{MO,ia}^x - \mathbf{S}_{MO,ia}^x \boldsymbol{\epsilon}}{\epsilon_a^{(0)} - \epsilon_i^{(0)}}. \quad (3.64)$$

While the virt/virt part of \mathbf{U}^x still vanishes the occ/occ part is completely determined by the derivative of the metric

$$\mathbf{U}_{ij}^x = -\mathbf{S}_{MO,ij}^x, \quad (3.65)$$

as can be seen from the derivative of the orthonormality condition ($\mathbf{C}^\dagger \mathbf{S} \mathbf{C} = \mathbf{1}$).

3.6 Linear Scaling Methods for Static Second Order Properties: First Order Density Matrix-based CP-SCF

As it is shown in section 2.5.3, the number of significant elements in the orbital coefficients matrix \mathbf{C} scales quadratically with system size. Therefore the previously described CPSCF schemes scale with $\mathcal{O}(M^3)$ for a single perturbation because of the AO-MO transformations ($\mathbf{A}_{AO} \rightarrow \mathbf{A}^{MO} = \mathbf{C}^\dagger \mathbf{A}^{AO} \mathbf{C}$). In the following the reformulation of the CPSCF equations entirely in the AO basis is presented which overcomes this last obstacle.

3.6.1 Linear Response Equations to a Static Perturbation

In the original work of Ochsenfeld and Head-Gordon [11] the D-CPSCF equations were derived from a quadratically convergent, density matrix-based SCF scheme (D-QCSCF). After a brief description of this diagonalization-free SCF alternative and the original D-CPSCF equations, a new formulation is presented which has been improved for the linear scaling calculation of molecular properties.

3.6.1.1 A Brief Review of the Quadratically Convergent Density Matrix-based SCF Energy Minimization

In this section we will derive a density matrix-based quadratically convergent SCF scheme (D-QCSCF) proposed by Ochsenfeld and Head-Gordon [11].

Ignoring the constant nuclear repulsion energy the electronic Hartree-Fock energy in terms of the density matrix \mathbf{P} is

$$E = \text{Tr} \left[\mathbf{P}\mathbf{h} + \frac{1}{2}\mathbf{P}\mathbf{G}(\mathbf{P}) \right]. \quad (3.66)$$

The following discussion is also valid for Kohn-Sham-DFT by just replacing the corresponding expressions in $\mathbf{G}[\mathbf{P}]$. In the pioneering work of Li, Nunes, and Vanderbilt (LNV) [48,87] the authors proposed an energy minimization procedure constrained to a fixed electron number N_{el} at fixed chemical potential μ

$$\mathcal{L}_{LNV} = E - \mu N_{el} = \text{Tr} \left[\tilde{\mathbf{P}}(\mathbf{h} + \frac{1}{2}\mathbf{G}(\mathbf{P}) - \mu) \right], \quad (3.67)$$

while ensuring the idempotency by the purification transformation of McWeeny in eq. 2.51. Note that eq. 3.67 contains the HF Hamiltonian while the original work [48] is concerned with the tight-binding approximation only. In contrast, we assume in our formulation the initial density to be in the quadratic basin of the stationary point, so that the number of electrons will be imposed by the purification transformation itself and an unconstrained minimization becomes possible. Inserting this purification transformation of eq. 2.51 into the energy expression eq. 3.66 we define the Lagrangian

$$\mathcal{L} = \text{Tr} \left[\tilde{\mathbf{P}}\mathbf{h} + \frac{1}{2}\tilde{\mathbf{P}}\mathbf{G}(\tilde{\mathbf{P}}) \right]. \quad (3.68)$$

In order to derive a working formula we expand the Lagrangian in \mathbf{P} around the minimum \mathbf{P}_{min}

$$\mathcal{L}(\mathbf{P}_{new}) = \mathcal{L}(\mathbf{P}) + \left. \frac{\partial \mathcal{L}(\mathbf{P})}{\partial \mathbf{P}} \right|_{\mathbf{P}_{\Delta}=0} \mathbf{P}_{\Delta} + \frac{1}{2} \left. \frac{\partial^2 \mathcal{L}(\mathbf{P})}{\partial \mathbf{P}^2} \right|_{\mathbf{P}_{\Delta}=0} (\mathbf{P}_{\Delta})^2 + \dots, \quad (3.69)$$

with $\mathbf{P}_{\Delta} = \mathbf{P}_{new} - \mathbf{P}$. Assuming to be close to the minimum ($\mathbf{P} = \mathbf{P}_{min}$) we can approximately use a quadratic form and truncate the expansion after the third term. Since the gradient at a stationary point vanishes, we obtain for $\mathbf{P} = \mathbf{P}_{min}$

$$\mathcal{L}(\mathbf{P}_{new}) = \mathcal{L}(\mathbf{P}_{min}) + \frac{1}{2} (\mathbf{P}_{\Delta})^2 \frac{\partial^2 \mathcal{L}(\mathbf{P}_{min})}{\partial \mathbf{P}_{min}^2}. \quad (3.70)$$

From the derivative of eq. 3.70 we also have

$$\frac{\partial \mathcal{L}(\mathbf{P}_{new})}{\partial \mathbf{P}_{new}} = \frac{\partial \mathcal{L}(\mathbf{P})}{\partial \mathbf{P}} + (\mathbf{P}_{\Delta}) \frac{\partial^2 \mathcal{L}(\mathbf{P})}{\partial \mathbf{P}^2}. \quad (3.71)$$

For $\mathbf{P}_{new} = \mathbf{P}_{min}$, i.e. "one step from minimum", we get from eq. 3.71

$$\left[\frac{\partial^2 \mathcal{L}(\mathbf{P})}{\partial \mathbf{P}^2} \right] \mathbf{P}_{\Delta} = - \frac{\partial \mathcal{L}(\mathbf{P})}{\partial \mathbf{P}}. \quad (3.72)$$

Note that the Hessian has to be non-singular^{a)} for the determination of \mathbf{P}_{Δ} .

Eq. 3.72 provides us with a working formula to optimize the density in an unconstrained fashion, as long as the initial guess $\mathbf{P}_{initial}$ is within the quadratic basin.

Using general rules for derivatives of traces^{b)} the gradient of the Lagrangian on the right-hand side (RHS) of eq. 3.72 is written as

$$\frac{\partial \mathcal{L}(\mathbf{P})}{\partial \mathbf{P}} = 3\mathbf{FPS} + 3\mathbf{SPF} - 2\mathbf{FPSPS} - 2\mathbf{SPFPS} - 2\mathbf{SPSPF}. \quad (3.74)$$

with \mathbf{F} as the Hartree-Fock or Kohn-Sham matrix

$$\mathbf{F} = \mathbf{h} + \mathbf{G}(\mathbf{P}). \quad (3.75)$$

Note that at convergence, when the gradient has to be equal to zero, eq. 3.74 reduces to $\mathbf{0} = \mathbf{FPS} - \mathbf{SPF}$. Within the optimization process the gradient in eq. 3.74 is used as

^{a)}Following the stability condition [88] a valid Hessian has to be positive-definite.

^{b)}Derivative of traces of n th order in \mathbf{A} :

$$\frac{\partial}{\partial A} \text{Tr}[\mathbf{BA}^n] = \sum_{m=0}^{n-1} (\mathbf{A}^m \mathbf{B} \mathbf{A}^{n-m-1})^\dagger \quad (3.73)$$

direction for the density matrix \mathbf{P} , thus the correct metric has of course to be accounted for. Head-Gordon et al. [89] have shown that the metric, i.e. the transformation behavior of the quantities, has to be considered explicitly. Since the density is contravariant by definition, the covariant gradient in eq. 3.74 has to be transformed to its contravariant representation by multiplication with the inverse metric \mathbf{S}^{-1}

$$\mathbf{S}^{-1} \frac{\partial \mathcal{L}(\mathbf{P})}{\partial \mathbf{P}} \mathbf{S}^{-1} = 3\mathbf{S}^{-1}\mathbf{F}\mathbf{P} + 3\mathbf{P}\mathbf{F}\mathbf{S}^{-1} - 2\mathbf{S}^{-1}\mathbf{F}\mathbf{P}\mathbf{S}\mathbf{P} - 2\mathbf{P}\mathbf{F}\mathbf{P} - 2\mathbf{P}\mathbf{S}\mathbf{P}\mathbf{F}\mathbf{S}^{-1}. \quad (3.76)$$

See e.g. Ref. [90] for an introduction to tensor theory in the framework of quantum chemistry. For a Newton-Raphson scheme the corresponding second derivative is

$$\begin{aligned} \frac{\partial}{\partial \mathbf{P}} \text{Tr} \left[\frac{\partial \mathcal{L}(\mathbf{P})}{\partial \mathbf{P}} \mathbf{P}_\Delta \right] &= 3\mathbf{F}\mathbf{P}_\Delta \mathbf{S} + 3\mathbf{S}\mathbf{P}_\Delta \mathbf{F} - 2\mathbf{F}\mathbf{P}_\Delta \mathbf{S}\mathbf{P}\mathbf{S} - 2\mathbf{F}\mathbf{P}\mathbf{S}\mathbf{P}_\Delta \mathbf{S} \\ &\quad - 2\mathbf{S}\mathbf{P}_\Delta \mathbf{F}\mathbf{P}\mathbf{S} - 2\mathbf{S}\mathbf{P}\mathbf{F}\mathbf{P}_\Delta \mathbf{S} - 2\mathbf{S}\mathbf{P}_\Delta \mathbf{S}\mathbf{P}\mathbf{F} - 2\mathbf{S}\mathbf{P}\mathbf{S}\mathbf{P}_\Delta \mathbf{F} \\ &\quad + 3\mathbf{G}(\mathbf{X})\mathbf{P}\mathbf{S} + 3\mathbf{S}\mathbf{P}\mathbf{G}(\mathbf{X}) - 2\mathbf{G}(\mathbf{X})\mathbf{P}\mathbf{S}\mathbf{P}\mathbf{S} \\ &\quad - 2\mathbf{S}\mathbf{P}\mathbf{S}\mathbf{P}\mathbf{G}(\mathbf{X}) - 2\mathbf{S}\mathbf{P}\mathbf{G}(\mathbf{X})\mathbf{P}\mathbf{S}, \end{aligned} \quad (3.77)$$

with

$$\mathbf{X} = 3\mathbf{P}_\Delta \mathbf{S}\mathbf{P} + 3\mathbf{P}\mathbf{S}\mathbf{P}_\Delta - 2\mathbf{P}_\Delta \mathbf{S}\mathbf{P}\mathbf{S}\mathbf{P} - 2\mathbf{P}\mathbf{S}\mathbf{P}_\Delta \mathbf{S}\mathbf{P} - 2\mathbf{P}\mathbf{S}\mathbf{P}\mathbf{S}\mathbf{P}_\Delta. \quad (3.78)$$

Inserting these expressions into eq. 3.72 we obtain a linear equation system (LEQS) that can be solved with standard numerical methods like the conjugate gradient algorithm.

All matrices that occur in eq. 3.77 are asymptotically linear scaling with respect to the number of significant elements for systems with a non-vanishing band gap, thus their sparsity can be exploited using sparse algebra routines (see chapter 4). In combination with the LinK and CFMM methods that have been discussed in the foregoing sections, an asymptotically overall linear scaling behavior is possible.

3.6.1.2 D-CPSCF Scheme Derived from D-QCSCF

As described before, the perturbation induces changes in the electron distribution and – as long as these are weak – an expansion like eq. 3.69 holds. Thus self-consistency can be maintained to first order within a linear response formalism, i.e. the response of the electronic structure to the external perturbation in first order,

$$\left[\frac{\partial^2 \mathcal{L}(\mathbf{P})}{\partial \mathbf{P}^2} \right] \mathbf{P}^x = -\frac{\partial}{\partial x} \left[\frac{\partial \mathcal{L}(\mathbf{P})}{\partial \mathbf{P}} \right], \quad (3.79)$$

with \mathbf{P}^x as a guess to the perturbed density matrix. Since the Hessian is known from eq. 3.77 with the replacement of \mathbf{P}_Δ by \mathbf{P}^x , we only have to consider the RHS. Note that for a converged density matrix all relations in section 2.5.2 hold, so that we can simplify the equation using the idempotency condition and the commutator with the Hamiltonian.

$$\begin{aligned} \frac{\partial}{\partial x} \left[\frac{\partial \mathcal{L}(\mathbf{P})}{\partial \mathbf{P}} \right] &= -\mathbf{F}\mathbf{P}^x\mathbf{S} - \mathbf{S}^x\mathbf{P}\mathbf{F} - 2\mathbf{F}\mathbf{P}\mathbf{S}^x\mathbf{P}\mathbf{S} - 2\mathbf{S}\mathbf{P}\mathbf{S}^x\mathbf{P}\mathbf{F} \\ &\quad -\mathbf{F}^x\mathbf{P}\mathbf{S} - \mathbf{S}\mathbf{P}\mathbf{F}^x + 2\mathbf{S}\mathbf{P}\mathbf{F}^x\mathbf{P}\mathbf{S}, \end{aligned} \quad (3.80)$$

with

$$\mathbf{F}^x = \mathbf{h}^x + \mathbf{G}^x(\mathbf{P}) + \mathbf{G}(\mathbf{P}^x). \quad (3.81)$$

In the original work [11] the term $\mathbf{G}(\mathbf{P}^x)$ is further split into the different parts of \mathbf{P}^x

$$\begin{aligned} \mathbf{G}(\mathbf{P}^x) &= \mathbf{G}(\mathbf{P}_{\text{ov}}^x + \mathbf{P}_{\text{vo}}^x) + \mathbf{G}(\mathbf{P}\mathbf{S}^x\mathbf{P}) \\ &= \mathbf{G}(\mathbf{P}\mathbf{S}\mathbf{P}^x + \mathbf{P}^x\mathbf{S}\mathbf{P} - 2\mathbf{P}\mathbf{S}\mathbf{P}^x\mathbf{S}\mathbf{P}) + \mathbf{G}(\mathbf{P}\mathbf{S}^x\mathbf{P}), \end{aligned} \quad (3.82)$$

so our preliminary D-CPSCF equations result as

$$\begin{aligned} &3\mathbf{F}\mathbf{P}^x\mathbf{S} + 3\mathbf{S}\mathbf{P}^x\mathbf{F} - 2\mathbf{F}\mathbf{P}^x\mathbf{S}\mathbf{P}\mathbf{S} - 4\mathbf{F}\mathbf{P}\mathbf{S}\mathbf{P}^x\mathbf{S} \\ &-4\mathbf{S}\mathbf{P}^x\mathbf{S}\mathbf{P}\mathbf{F} - 2\mathbf{S}\mathbf{P}\mathbf{S}\mathbf{P}^x\mathbf{F} + \mathbf{G}(\mathbf{X}')\mathbf{P}\mathbf{S} \\ &+\mathbf{S}\mathbf{P}\mathbf{G}(\mathbf{X}') - 2\mathbf{S}\mathbf{P}\mathbf{G}(\mathbf{X}')\mathbf{P}\mathbf{S} \\ &= \mathbf{F}\mathbf{P}\mathbf{S}^x + \mathbf{S}^x\mathbf{P}\mathbf{F} + 2\mathbf{F}\mathbf{P}\mathbf{S}^x\mathbf{P}\mathbf{S} + 2\mathbf{S}\mathbf{P}\mathbf{S}^x\mathbf{P}\mathbf{F} \\ &\quad -\mathbf{F}^{(x)'}\mathbf{P}\mathbf{S} - \mathbf{S}\mathbf{P}\mathbf{F}^{(x)'} + 2\mathbf{S}\mathbf{P}\mathbf{F}^{(x)'}\mathbf{P}\mathbf{S}, \end{aligned} \quad (3.83)$$

with

$$\mathbf{F}^{(x)'} = \mathbf{h}^x + \mathbf{G}^x(\mathbf{P}) - \mathbf{G}(\mathbf{P}\mathbf{S}^x\mathbf{P}), \quad (3.84)$$

$$\begin{aligned} \mathbf{X}' &= \mathbf{P}^x\mathbf{S}\mathbf{P} + \mathbf{P}\mathbf{S}\mathbf{P}^x - 2\mathbf{P}\mathbf{S}\mathbf{P}^x\mathbf{S}\mathbf{P} \\ &= \mathbf{P}_{\text{ov}}^x + \mathbf{P}_{\text{vo}}^x. \end{aligned} \quad (3.85)$$

In order to get a more convenient expression for the Fock-builds $\mathbf{F}^{(x)'}$ and $\mathbf{G}(\mathbf{X}')$ we define

$$\mathbf{F}^{(x)} = \mathbf{h}^x + \mathbf{G}^x(\mathbf{P}), \quad (3.86)$$

$$\begin{aligned} \mathbf{X} &= \mathbf{P}^x\mathbf{S}\mathbf{P} + \mathbf{P}\mathbf{S}\mathbf{P}^x - 2\mathbf{P}\mathbf{S}\mathbf{P}^x\mathbf{S}\mathbf{P} - \mathbf{P}\mathbf{S}^x\mathbf{P} \\ &= \mathbf{P}^x, \end{aligned} \quad (3.87)$$

which replace the quantities $\mathbf{F}^{(x)'}$ and \mathbf{X}' . Note in eq. 3.87 the distinction between the projection $\mathbf{P}_{\text{oo}}^x = \mathbf{PSP}^x\mathbf{SP}$ and $-\mathbf{PS}^x\mathbf{P}$, which are in principle equal as seen in eq. 3.13. This is necessary since the projections of the intermediate quantity \mathbf{P}^x may not obey the equality.

The LEQS in eq. 3.83 provides a basis for a CPSCF calculation, but within an iterative algorithm instabilities may occur caused by approximate quantities, in particular if sparse algebra routines are employed. Consider for example an initial guess $\mathbf{P}^{x'}$ for the perturbed density matrix, that has a non-vanishing virt/virt part $\mathbf{P}_{\text{vv}}^{x'}$. It is easily seen that $\mathbf{P}_{\text{vv}}^{x'}$ vanishes if it is multiplied with \mathbf{P}_{occ} because of the orthogonality of the subspaces (eq. 2.48):

$$\mathbf{P}_{\text{occ}}\mathbf{P}_{\text{vv}}^{x'} = \mathbf{P}_{\text{occ}}\mathbf{P}_{\text{virt}}\mathbf{P}^{x'}\mathbf{P}_{\text{virt}} = \mathbf{0}. \quad (3.88)$$

The drawback in eq. 3.83 is found in the first two terms that remain "unprojected":

$$\mathbf{FP}_{\text{vv}}^{x'}\mathbf{S} \neq \mathbf{0}, \quad \mathbf{SP}_{\text{vv}}^{x'}\mathbf{F} \neq \mathbf{0}. \quad (3.89)$$

Thus, the virt/virt part $\mathbf{P}_{\text{vv}}^{x'}$ would not vanish even in an optimization process where all quantities are treated with high accuracy, i.e. no sparse algebra and tight integral thresholds are applied. For similar reasons one can also see that a non-valid guess to the occ/occ part $\mathbf{P}_{\text{oo}}^{x'} \neq -\mathbf{PS}^x\mathbf{P}$ would sustain. It has to be mentioned that these problems become more severe if additionally sparse algebra routines are applied. Thus the method has to be improved by imposing the constraints on \mathbf{P}^x in the LEQS itself by using the projection properties of \mathbf{P} . The occ/occ part \mathbf{P}_{oo}^x can be determined at the beginning of the calculation and so provides an initial guess to \mathbf{P}^x . Since the \mathbf{P}_{vv}^x has to vanish, only the occ/virt and virt/occ parts have to be determined. Projecting eq. 3.83 from the left with \mathbf{SP} and from the right with \mathbf{PS} yields the following two equations:

$$\begin{aligned} \mathbf{SPSP}^x\mathbf{F} - \mathbf{FPSP}^x\mathbf{S} + \mathbf{SPG}(\mathbf{X}) - \mathbf{SPG}(\mathbf{X})\mathbf{PS} \\ = \mathbf{FPS}^x - \mathbf{SPS}^x\mathbf{PF} - \mathbf{SPF}^{(x)} + \mathbf{SPF}^{(x)}\mathbf{PS}, \end{aligned} \quad (3.90)$$

$$\begin{aligned} \mathbf{FP}^x\mathbf{SPS} - \mathbf{SP}^x\mathbf{SPF} + \mathbf{G}(\mathbf{X})\mathbf{PS} - \mathbf{SPG}(\mathbf{X})\mathbf{PS} \\ = \mathbf{S}^x\mathbf{PF} - \mathbf{FPS}^x\mathbf{PS} - \mathbf{F}^{(x)}\mathbf{PS} + \mathbf{SPF}^{(x)}\mathbf{PS}. \end{aligned} \quad (3.91)$$

The new LEQS is constructed as $(\mathbf{SP} (\text{LEQS}_{eq. 3.83}) + (\text{LEQS}_{eq. 3.83}) \mathbf{PS}) = \text{LEQS}_{eq. 3.90} +$

LEQS_{eq. 3.91})

$$\begin{aligned}
 & \mathbf{FP}^x\mathbf{SPS} - \mathbf{FPSP}^x\mathbf{S} + \mathbf{SPSP}^x\mathbf{F} - \mathbf{SP}^x\mathbf{SPF} \\
 & + \mathbf{G}(\mathbf{X})\mathbf{PS} + \mathbf{SPG}(\mathbf{X}) - 2\mathbf{SPG}(\mathbf{X})\mathbf{PS} \\
 = & \mathbf{FPS}^x + \mathbf{S}^x\mathbf{PF} - \mathbf{SPS}^x\mathbf{PF} - \mathbf{FPS}^x\mathbf{PS} \\
 & - \mathbf{F}^{(x)}\mathbf{PS} - \mathbf{SPF}^{(x)} + 2\mathbf{SPF}^{(x)}\mathbf{PS}.
 \end{aligned} \tag{3.92}$$

Obviously, the perturbed density matrix in \mathbf{X} is inert to projections of the LEQS, but it is "cleaned" by eq. 3.87, i.e. the redundant parts \mathbf{P}_{oo}^x and \mathbf{P}_{vv}^x are removed by subtraction.

Since \mathbf{P}^x will be updated by the residual ($RHS - LHS$) it is also necessary that the LEQS reassembles the essential parts of the perturbed density, i.e. that all redundant or a priori known parts are eliminated:

$$\begin{aligned}
 \text{LEQS}_{eq. 3.92} &= \mathbf{SP} (\text{LEQS}_{eq. 3.83}) + (\text{LEQS}_{eq. 3.83}) \mathbf{PS} \\
 &= (\text{LEQS}_{eq. 3.83})_{ov} + (\text{LEQS}_{eq. 3.83})_{vo}.
 \end{aligned} \tag{3.93}$$

Inserting the different projections of \mathbf{P}^x into the first four terms of the LHS of eq. 3.92 yields:

$$\begin{aligned}
 \mathbf{P}_{oo}^x : & \quad \mathbf{FP}_{oo}^x\mathbf{SPS} - \mathbf{FPSP}_{oo}^x\mathbf{S} + \mathbf{SPSP}_{oo}^x\mathbf{F} - \mathbf{SP}_{oo}^x\mathbf{SPF} = \mathbf{0}, \\
 \mathbf{P}_{ov}^x : & \quad \mathbf{FP}_{ov}^x\mathbf{SPS} - \mathbf{FPSP}_{ov}^x\mathbf{S} + \mathbf{SPSP}_{ov}^x\mathbf{F} - \mathbf{SP}_{ov}^x\mathbf{SPF} = \mathbf{F}(-\mathbf{P}_{ov}^x)\mathbf{S} - \mathbf{S}(-\mathbf{P}_{ov}^x)\mathbf{F}, \\
 \mathbf{P}_{vo}^x : & \quad \mathbf{FP}_{vo}^x\mathbf{SPS} - \mathbf{FPSP}_{vo}^x\mathbf{S} + \mathbf{SPSP}_{vo}^x\mathbf{F} - \mathbf{SP}_{vo}^x\mathbf{SPF} = \mathbf{FP}_{vo}^x\mathbf{S} - \mathbf{SP}_{vo}^x\mathbf{F}, \\
 \mathbf{P}_{vv}^x : & \quad \mathbf{FP}_{vv}^x\mathbf{SPS} - \mathbf{FPSP}_{vv}^x\mathbf{S} + \mathbf{SPSP}_{vv}^x\mathbf{F} - \mathbf{SP}_{vv}^x\mathbf{SPF} = \mathbf{0}.
 \end{aligned}$$

The last three terms of the LHS also form a "cleaned" representation of $\mathbf{G}(\mathbf{P}^x)$ by construction of the LEQS

$$\mathbf{G}(\mathbf{X})\mathbf{PS} + \mathbf{SPG}(\mathbf{X}) - 2\mathbf{SPG}(\mathbf{X})\mathbf{PS} = \mathbf{G}(\mathbf{P}^x)_{ov} + \mathbf{G}(\mathbf{P}^x)_{vo}.$$

As test calculations have shown, even a reconstruction of the RHS by reprojecting it onto the occ/virt and virt/occ subspaces is essential to ensure convergence of the linear equation solver, as it is most evident with a less tight integral threshold.

The advantage of using the difference between \mathbf{P}_{vo}^x and \mathbf{P}_{ov}^x in eq. 3.92 will be shown in the next section. As we will see later, it becomes even crucial when sparse algebra is applied.

3.6.1.3 An Alternative Derivation of D-CPSCF

In order to elucidate the advantages of our LEQS compared to similar CPSCF schemes like the one in Ref. [91] we will derive the equation starting from the commutator in eq. 2.46.

Differentiation of eq. 2.46 with respect to the perturbation x yields

$$\frac{\partial}{\partial x} (\mathbf{FPS} - \mathbf{SPF}) = \mathbf{F}^x \mathbf{PS} + \mathbf{FP}^x \mathbf{S} + \mathbf{FPS}^x - \mathbf{S}^x \mathbf{PF} - \mathbf{SP}^x \mathbf{F} - \mathbf{SPF}^x = \mathbf{0}. \quad (3.94)$$

Projecting from left with \mathbf{SP} and from right with \mathbf{PS} gives

$$\mathbf{SPF}^x \mathbf{PS} + \mathbf{FPSP}^x \mathbf{S} + \mathbf{FPS}^x - \mathbf{SPS}^x \mathbf{PF} - \mathbf{SPSP}^x \mathbf{F} - \mathbf{SPF}^x = \mathbf{0}, \quad (3.95)$$

$$\mathbf{F}^x \mathbf{PS} + \mathbf{FP}^x \mathbf{SPS} + \mathbf{FPS}^x \mathbf{PS} - \mathbf{S}^x \mathbf{PF} - \mathbf{SP}^x \mathbf{SPF} - \mathbf{SPF}^x \mathbf{PS} = \mathbf{0}, \quad (3.96)$$

where we have already employed the commutator $\mathbf{FPS} = \mathbf{SPF}$ and the idempotency of \mathbf{P} . By subtracting eq. 3.95 from eq. 3.96 we obtain $((\text{LEQS}_{eq. 3.94}) \mathbf{PS} - \mathbf{SP} (\text{LEQS}_{eq. 3.94}) = \text{LEQS}_{eq. 3.96} - \text{LEQS}_{eq. 3.95})$

$$\begin{aligned} & \mathbf{FP}^x \mathbf{SPS} + \mathbf{SPSP}^x \mathbf{F} - \mathbf{SP}^x \mathbf{SPF} - \mathbf{FPSP}^x \mathbf{S} + \mathbf{FPS}^x \mathbf{PS} \\ & + \mathbf{SPS}^x \mathbf{PF} - \mathbf{FPS}^x - \mathbf{S}^x \mathbf{PF} + \mathbf{F}^x \mathbf{PS} + \mathbf{SPF}^x - 2\mathbf{SPF}^x \mathbf{PS} = \mathbf{0}, \end{aligned} \quad (3.97)$$

and after shifting the terms without \mathbf{P}^x to the RHS we obtain eq. 3.92 again

$$\begin{aligned} & \mathbf{FP}^x \mathbf{SPS} - \mathbf{FPSP}^x \mathbf{S} + \mathbf{SPSP}^x \mathbf{F} - \mathbf{SP}^x \mathbf{SPF} \\ & + \mathbf{G}(\mathbf{P}^x) \mathbf{PS} + \mathbf{SPG}(\mathbf{P}^x) - 2\mathbf{SPG}(\mathbf{P}^x) \mathbf{PS} \\ = & \mathbf{FPS}^x + \mathbf{S}^x \mathbf{PF} - \mathbf{SPS}^x \mathbf{PF} - \mathbf{FPS}^x \mathbf{PS} \\ & - \mathbf{F}^{(x)} \mathbf{PS} - \mathbf{SPF}^{(x)} + 2\mathbf{SPF}^{(x)} \mathbf{PS}. \end{aligned} \quad (3.98)$$

As it is shown in the succeeding section, the LEQS proposed by Larsen et al. [91] can be constructed by addition of eq. 3.95 and eq. 3.96. The advantage of a subtraction used in our approach is the higher numerical stability of eq. 3.98 resulting from a cancellation of numerical inaccuracies. This is of course particularly important when sparse algebra routines are used.

Considering the resolution in eqs. 3.13-3.16 one can see that the explicit formula of \mathbf{PSP}^x contains the difference between two occ/occ-projections of \mathbf{P}^x :

$$\mathbf{PSP}^x = \mathbf{P}_{\text{ov}}^x + \mathbf{P}_{\text{oo}}^x = \underbrace{\mathbf{PSP}^x - \mathbf{P}_{\text{oo}}^{x'}}_{\mathbf{P}_{\text{ov}}^x} + \mathbf{P}_{\text{oo}}^x. \quad (3.99)$$

If exact matrix-algebra routines are applied, the last two terms cancel each other, but within numerical accuracy $\mathbf{P}_{oo}^{x'}$ (occ/occ part of $\mathbf{P}_{ov}^x/\mathbf{P}_{vo}^x$) and \mathbf{P}_{oo}^x (initial guess: $-\mathbf{P}\mathbf{S}^x\mathbf{P}$) may differ and so produce numerical noise (see Tab. 4.2 and the discussion in chapter 4)

$$\mathbf{PSP}^x = \mathbf{PSP}^x + \Delta\mathbf{P}_{oo}^x. \quad (3.100)$$

From the symmetry $\mathbf{P}_{ov}^x = \mathbf{P}_{vo}^{x\dagger}$ we see that both terms would vanish by subtraction

$$\begin{aligned} \mathbf{P}^x\mathbf{S}\mathbf{P} - \mathbf{PSP}^x &= \mathbf{P}_{vo}^x + \mathbf{P}_{oo}^x - \mathbf{P}_{ov}^x - \mathbf{P}_{oo}^x \\ &= \underbrace{\mathbf{P}^x\mathbf{S}\mathbf{P} - \mathbf{P}_{oo}^{x'}}_{\mathbf{P}_{vo}^x} - \underbrace{\mathbf{PSP}^x + \mathbf{P}_{oo}^{x'}}_{\mathbf{P}_{ov}^x} \\ &= \mathbf{P}^x\mathbf{S}\mathbf{P} - \mathbf{PSP}^x, \end{aligned} \quad (3.101)$$

where the equality of $\mathbf{P}_{oo}^{x'}$ from \mathbf{P}_{ov}^x and \mathbf{P}_{vo}^x is given by the hermiticity of \mathbf{P}^x ensured by the symmetry of the LEQS. Since the perturbed density is cleaned by the form of the LEQS, i.e. the subtraction is intrinsic to eq. 3.92, we can also skip the explicit reprojection of \mathbf{P}^x in \mathbf{X} (eq. 3.87).

Note that within an iterative scheme the numerical noise will increase – comparable to an ”autocatalytic” process – which will not only affect the results but as well decrease the sparsity of the matrices crucial for the efficient application of sparse algebra routines. See chapter 4 for a more detailed analysis and implementational details.

Applying sparse-matrix algebra we have observed a distinct decrease in the sparsity of the perturbed densities^{c)} resulting from the truncation of the matrices at small values. This originates from the transformation of the residual matrix \mathbf{R} from the co- to the contravariant basis within the conjugate gradient routine in order to be consistent with tensor theory (e.g. Ref. [90])

$$\Delta\mathbf{P} = \mathbf{S}^{-1}\mathbf{R}\mathbf{S}^{-1}. \quad (3.102)$$

Since this transformation is essential for the convergence of the linear equation solver, we have to reduce the number of multiplications with the inverse metric which is less sparse than any other matrix in eq. 3.98. A first step is to reformulate the transformation as a combination of transformation and projection onto the ov-vo subspace

$$\mathbf{P}^x = \mathbf{P}\mathbf{R}\mathbf{S}^{-1} + \mathbf{S}^{-1}\mathbf{R}\mathbf{P}. \quad (3.103)$$

^{c)}The effect does not occur when using minimal basis sets that have been used in other publications, e.g. Ref. [15].

As we have observed from a projection of the intermediate \mathbf{P}^x matrices onto the vv-subspace, the numerical noise produced from the application of the inverse metric enters mainly \mathbf{P}_{vv}^x . Using formula 3.103, the decrease is significantly smaller since we implicitly project onto the ov-vo-subspace and so delete the vv-part.

In section 4.3 it is shown that multiplies with the inverse metric \mathbf{S}^{-1} exhibit the largest error in combination with sparse algebra routines. This results from the large condition number of \mathbf{S}^{-1} , i.e. large absolute values of the elements in \mathbf{S}^{-1} compared to the other matrices. Challacombe [14] proposed to circumvent the explicit application of the inverse by multiplications with the inverse Cholesky factor that exhibits a smaller condition number and a more sparse structure. In this work the explicit use of \mathbf{S}^{-1} is minimized with the intermediate

$$\tilde{\mathbf{F}} = \mathbf{S}^{-1}\mathbf{F} \quad (3.104)$$

that is built only once. Additionally the whole equation is transformed into the contravariant basis:

$$\begin{aligned} & \tilde{\mathbf{F}}\mathbf{P}^x\mathbf{S}\mathbf{P} + \mathbf{P}\mathbf{S}\mathbf{P}^x\tilde{\mathbf{F}}^\dagger - \mathbf{P}\mathbf{F}\mathbf{P}^x - \mathbf{P}^x\mathbf{F}\mathbf{P} \\ & + \mathbf{S}^{-1}\mathbf{G}(\mathbf{P}^x)\mathbf{P} + \mathbf{P}\mathbf{G}(\mathbf{P}^x)\mathbf{S}^{-1} - 2\mathbf{P}\mathbf{G}(\mathbf{P}^x)\mathbf{P} \\ = & \tilde{\mathbf{F}}\mathbf{P}\mathbf{S}^x\mathbf{S}^{-1} + \mathbf{S}^{-1}\mathbf{S}^x\mathbf{P}\tilde{\mathbf{F}}^\dagger - \tilde{\mathbf{F}}\mathbf{P}\mathbf{S}^x\mathbf{P} - \mathbf{P}\mathbf{S}^x\mathbf{P}\tilde{\mathbf{F}}^\dagger \\ & - \mathbf{S}^{-1}\mathbf{Y}\mathbf{P} - \mathbf{P}\mathbf{Y}\mathbf{S}^{-1} + 2\mathbf{P}\mathbf{Y}\mathbf{P}. \end{aligned} \quad (3.105)$$

At a first glance the use of this equation seems to involve a larger number of operations with the inverse metric \mathbf{S}^{-1} compared to eq. 3.98. If we consider the implementation of the D-CPSCF algorithm described in section 3.6.3.1, we see that the RHS is only built once and the last three terms of the LHS are only formed at Level-2. The first four terms are iterated within a conjugate gradient algorithm where the effect of avoiding \mathbf{S}^{-1} in eq. 3.103 and the substitution with $\tilde{\mathbf{F}}$ is most important.

This equation combined with the use of an extrapolated initial residual matrix for the conjugate gradient routine is found to be most efficient to prevent a decrease of sparsity in the occurring quantities.

3.6.1.4 Comparison with Other CPSCF Algorithms

To derive the conventional approach of Frisch et al. [92] we have to project out the occ/virt part of eq. 3.98, change signs and switch to an orthonormal basis ($\mathbf{S} = \mathbf{1}$)

$$\mathbf{F}\mathbf{P}_{\text{ov}}^x - \mathbf{P}_{\text{ov}}^x\mathbf{F} + \mathbf{G}_{\text{ov}}(\mathbf{P}_{\text{ov}}^x + \mathbf{P}_{\text{vo}}^x) = \mathbf{F}_{\text{ov}}^{(x)} - \mathbf{G}_{\text{ov}}(\mathbf{S}_{\text{oo}}^x) - \mathbf{F}\mathbf{S}_{\text{ov}}^x. \quad (3.106)$$

In the original approach these equations are solved within the MO basis. As our test calculations have shown, the calculations performed in the AO basis exhibit a poor convergence behavior. The reason is that the constraints on the perturbed density \mathbf{P}^x are not considered by the LEQS in eq. 3.106. Therefore, the construction of \mathbf{P}^x from the first order MO coefficients becomes necessary which directly ensures the proper subspace projections of \mathbf{P}^x .

Another algorithm for the determination of the linear response totally within the AO basis has been proposed later by Larsen et al. [91]. Their formalism is based on an exponential parametrization of the one-electron density matrix [12, 13, 93] utilizing the asymmetric Baker-Campbell-Hausdorff expansion

$$\begin{aligned} \mathbf{P}(\Delta) &= e^{-\Delta\mathbf{S}}\mathbf{P}e^{\mathbf{S}\Delta} \\ &= \mathbf{P} + [\mathbf{P}, \Delta]_{\mathbf{S}} + \frac{1}{2} [[\mathbf{P}, \Delta]_{\mathbf{S}}, \mathbf{P}]_{\mathbf{S}} + \dots, \end{aligned} \quad (3.107)$$

with the \mathbf{S} commutator $[\mathbf{A}, \mathbf{B}]_{\mathbf{S}} = \mathbf{A}\mathbf{S}\mathbf{B} - \mathbf{B}\mathbf{S}\mathbf{A}$ and an anti-Hermitian matrix Δ that only implies non-redundant orbital rotations, i.e. only contains occ/virt and virt/occ parts

$$\Delta = \Delta_{\text{ov}} + \Delta_{\text{vo}}. \quad (3.108)$$

It has to be mentioned that this approach leads to the efficient "curvy-steps" method to minimize the ground-state density matrix \mathbf{P} [12, 13]. This approach shows an improved convergence behavior within a density matrix-based SCF algorithm if a less accurate guess to the ground-state density matrix \mathbf{P} is provided. For a CPSCF calculation, however, only a small deviation \mathbf{P}^x is assumed, thus the exponential parametrization in eq. 3.107 is actually redundant.

The resulting linear response equations of Larsen et al. [91] are

$$\begin{aligned} &\mathbf{G}([\mathbf{P}, \Delta^x]_{\mathbf{S}})\mathbf{P}\mathbf{S} + \mathbf{F}[\mathbf{P}, \Delta^x]_{\mathbf{S}}\mathbf{S} - \mathbf{S}\mathbf{P}\mathbf{G}([\mathbf{P}, \Delta^x]_{\mathbf{S}}) - \mathbf{S}[\mathbf{P}, \Delta^x]_{\mathbf{S}}\mathbf{F} \\ = &\mathbf{S}\mathbf{P}\mathbf{F}^{(x)} - \mathbf{F}^{(x)}\mathbf{P}\mathbf{S} - \mathbf{S}\mathbf{P}\mathbf{S}^x\mathbf{P}\mathbf{F} + \mathbf{F}\mathbf{P}\mathbf{S}^x\mathbf{P}\mathbf{S} + \mathbf{S}^x\mathbf{P}\mathbf{F} - \mathbf{F}\mathbf{P}\mathbf{S}^x \\ &- \mathbf{S}\mathbf{P}\mathbf{G}(\mathbf{P}\mathbf{S}^x\mathbf{P}) + \mathbf{G}(\mathbf{P}\mathbf{S}^x\mathbf{P})\mathbf{P}\mathbf{S}. \end{aligned} \quad (3.109)$$

Identifying Δ^x with $\mathbf{P}_{\text{ov}}^x - \mathbf{P}_{\text{vo}}^x$ we can switch to our usual notation

$$\begin{aligned} & \mathbf{F}\mathbf{P}_{\text{ov}}^x\mathbf{S} + \mathbf{F}\mathbf{P}_{\text{vo}}^x\mathbf{S} - \mathbf{S}\mathbf{P}_{\text{ov}}^x\mathbf{F} - \mathbf{S}\mathbf{P}_{\text{vo}}^x\mathbf{F} + \mathbf{G}(\mathbf{P}_{\text{ov}}^x + \mathbf{P}_{\text{vo}}^x)\mathbf{P}\mathbf{S} - \mathbf{S}\mathbf{P}\mathbf{G}(\mathbf{P}_{\text{ov}}^x + \mathbf{P}_{\text{vo}}^x) \\ = & \mathbf{S}\mathbf{P}\mathbf{F}^{(x)} - \mathbf{F}^{(x)}\mathbf{P}\mathbf{S} - \mathbf{S}\mathbf{P}\mathbf{S}^x\mathbf{P}\mathbf{F} + \mathbf{F}\mathbf{P}\mathbf{S}^x\mathbf{P}\mathbf{S} + \mathbf{S}^x\mathbf{P}\mathbf{F} - \mathbf{F}\mathbf{P}\mathbf{S}^x \\ & + \mathbf{G}(\mathbf{P}\mathbf{S}^x\mathbf{P})\mathbf{P}\mathbf{S} - \mathbf{S}\mathbf{P}\mathbf{G}(\mathbf{P}\mathbf{S}^x\mathbf{P}). \end{aligned} \quad (3.110)$$

A comparison with the derivation of our D-CPSCF equations (eq. 3.98) shows that the latter equation would be obtained by addition of eq. 3.95 and eq. 3.96 instead of the subtraction which we use. From the previous discussion it is obvious that the performance of this LEQS in combination with sparse algebra routines is hampered by the accumulation of numerical noise in the *occ/occ* and *virt/virt* parts of \mathbf{P}^x . Thus an efficient treatment with sparse algebra routines is less efficient as our test calculations have shown.

3.6.2 Nuclear Magnetic Shielding Tensor

Modern high-field Fourier transform nuclear magnetic resonance spectroscopy (NMR) is a powerful technique to explore many different types of problems in chemistry and biochemistry. Since there is no simple and direct relationship between the measured NMR signals and structural properties of the adjacent environment, the necessity for a reliable method to predict NMR chemical shifts for large systems arises.

In this section the fundamental equations [94, 95] for the ab initio treatment of NMR chemical shielding tensors are presented. The basic physical effect in NMR spectroscopy is the induction of an electronic current caused by the external magnetic field which results in an additional, induced field. The effective magnetic field \mathbf{B}_{eff} that is experienced by the nuclei is the sum of the external (\mathbf{B}) and the induced (\mathbf{B}_{ind}) field.

$$\mathbf{B}_{\text{eff}} = \mathbf{B} + \mathbf{B}_{\text{ind}}, \quad \mathbf{B}_{\text{ind}} = -\boldsymbol{\sigma}\mathbf{B}. \quad (3.111)$$

The proportionality constant between the external and the induced B-field is the magnetic shielding tensor^{d)}. The difference of the nuclear spin energy levels is

$$\Delta E = -\mathbf{m}_j(1 - \boldsymbol{\sigma})\mathbf{B}. \quad (3.112)$$

As the shielding tensor $\boldsymbol{\sigma}$ appears in a term bilinear in the external field \mathbf{B} and the nuclear magnetic spin moment \mathbf{m}_j of nucleus j , $\boldsymbol{\sigma}$ is a second order property relative to

^{d)}the induced field is not necessarily parallel to the external field.

the perturbations \mathbf{B} and \mathbf{m}_j . If we treat the external magnetic field and the field induced by the nuclear magnetic moments as weak perturbations, we can expand the energy in a Taylor series around the unperturbed state

$$E(\mathbf{B}, \mathbf{m}) = E(0, 0) + \left. \frac{dE}{d\mathbf{B}} \right|_{\mathbf{B}=0} \mathbf{B} + \left. \frac{dE}{d\mathbf{m}} \right|_{\mathbf{m}=0} \mathbf{m} + \left. \frac{d^2E}{d\mathbf{B}^2} \right|_{\mathbf{B}=0} \mathbf{B}^2 + \left. \frac{d^2E}{d\mathbf{m}^2} \right|_{\mathbf{m}=0} \mathbf{m}^2 + \left. \frac{d^2E}{d\mathbf{B}d\mathbf{m}} \right|_{\mathbf{B}, \mathbf{m}=0} \mathbf{B}\mathbf{m} + \dots \quad (3.113)$$

From Eq. 3.113 we can identify the last term in the Taylor expansion as the magnetic shielding tensor, i.e. it is calculated as second derivative of the electronic energy:

$$\boldsymbol{\sigma} = \left(\frac{d^2E}{d\mathbf{B} d\mathbf{m}_j} \right)_{\mathbf{B}, \mathbf{m}_j=0} \quad (3.114)$$

Starting from the expression for the energy functional

$$E = \text{Tr} \left(\mathbf{P}\mathbf{h} + \frac{1}{2}\mathbf{P}\mathbf{G}[\mathbf{P}] \right) \quad (3.115)$$

with $\mathbf{G}[\mathbf{P}] = \mathbf{\Pi}$ and \mathbf{h} as the matrix representation of 1-electron operators and $\mathbf{\Pi}$ as the 4-index matrix of the antisymmetrized 2-electron integrals (Mulliken notation):

$$G_{\mu\nu}[\mathbf{P}] = \sum_{\lambda\sigma} P_{\lambda\sigma} \left[(\phi_\mu\phi_\nu|\phi_\lambda\phi_\sigma) - \frac{1}{2}(\phi_\mu\phi_\sigma|\phi_\lambda\phi_\nu) \right]. \quad (3.116)$$

We use analytic techniques to evaluate the (3×3) shielding tensor $\boldsymbol{\sigma}$

$$\sigma_{\alpha\beta,j} = \sum_{\mu\nu} P_{\mu\nu} \frac{\partial^2 h_{\mu\nu}}{\partial B_\alpha \partial m_{\beta,j}} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial B_\alpha} \frac{\partial h_{\mu\nu}}{\partial m_{\beta,j}} \quad \text{with } \alpha, \beta = x, y, z. \quad (3.117)$$

3.6.2.1 Molecular Hamiltonian in the Presence of an External Magnetic Field — The Gauge-Origin Problem

In order to calculate the second derivatives we have to formulate the molecular Hamiltonian in the presence of a magnetic field. The charges experience an additional Lorentz force, i.e. a perturbation in the momenta occur, which is perpendicular to the magnetic field and the kinetic energy vector of the electron. Thus the canonical momentum \mathbf{p} is replaced by the kinetic momentum \mathbf{p}_B according to the principle of minimal electromagnetic coupling [96]:

$$\mathbf{p}_B = \mathbf{p} + \mathbf{A}, \quad (3.118)$$

where \mathbf{A} is the vector potential (eq. 3.119) describing not only the magnetic field (Eq. 3.120) but also the potential associated with the nuclear magnetic moments (Eq. 3.121) [96, 97].

$$\mathbf{A}(\mathbf{r}) = \mathbf{A}_B(\mathbf{r}) + \sum_j \mathbf{A}_j(\mathbf{r}), \quad (3.119)$$

$$\mathbf{A}_B(\mathbf{r}) = \frac{1}{2} \mathbf{B} \times (\mathbf{r} - \mathbf{R}_0), \quad (3.120)$$

$$\mathbf{A}_j(\mathbf{r}) = \alpha^2 \frac{\mathbf{m}_j \times (\mathbf{r} - \mathbf{R}_j)}{|\mathbf{r} - \mathbf{R}_j|^3}. \quad (3.121)$$

Here \mathbf{R}_0 is the gauge origin, \mathbf{R}_j the position of nucleus j and α the fine structure constant. Inserting eq. 3.118 into the Hamiltonian for a closed-shell system yields

$$\hat{H}(\mathbf{B}, \mathbf{m}) = \frac{1}{2} \sum_i \left[\hat{p} + \hat{A} \right]^2 - \sum_{i,A} \frac{Z_A}{r_{iA}} + \frac{1}{2} \sum_{i \neq j} \frac{1}{r_{ij}} + \frac{1}{2} \sum_{A \neq B} \frac{Z_A Z_B}{R_{AB}}. \quad (3.122)$$

To solve the Schrödinger equation of a system in presence of a magnetic field we have to find a wave function $\Psi(\mathbf{B})$ to the Hamiltonian in eq. 3.122. Without an external magnetic field the MO's $\phi_i(\mathbf{B})$ are expanded as a linear combination of real atomic orbitals. However, this approach leads to the gauge problem [79] if finite-size basis sets are used. While the vector potential \mathbf{A} is uniquely defined by the magnetic field

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad (3.123)$$

the opposite does not hold in eq. 3.120. One can add a gradient of an arbitrary scalar function ∇f to \mathbf{A} still obtaining the same result for \mathbf{B} since the curl of the gradient vanishes. This arbitrariness in the choice of the gauge-origin R_0 in eq. 3.120 does not affect the results obtained with wave functions represented in a complete basis [79, 98]. Apart from the need for a unique choice of the gauge origin the calculations also suffer from a slow convergence of the NMR shielding constants with respect to the basis set size (see figures in Ref. [79]).

Thus – according to London [99, 100] – a basis of gauge including atomic orbitals (GIAO) can be chosen:

$$\begin{aligned} \chi_\mu(\mathbf{B}) &= \chi_\mu(0) \exp \left(-\frac{i}{2} \mathbf{B} \times (\mathbf{R}_\mu - \mathbf{R}_0) \cdot \mathbf{r} \right) \\ &= \chi_\mu(0) f_\mu. \end{aligned} \quad (3.124)$$

The exponential gauge prefactor contains the field dependency of the basis function, i.e. it describes the B-field at the origin \mathbf{R}_μ of the function χ_μ . Because of its efficiency and reliability, this approach is nowadays the standard choice in most ab initio programs [79, 98, 101–103]. The wave function is built as a single determinant of the occupied MO's:

$$\Psi(\mathbf{B}) = |\phi_1(\mathbf{B}), \phi_2(\mathbf{B}), \dots, \phi_N(\mathbf{B})\rangle. \quad (3.125)$$

The energy $E(\mathbf{B}, \mathbf{m}_j)$ associated with $\Psi(\mathbf{B}, \mathbf{m}_j)$ is

$$E(\mathbf{B}, \mathbf{m}_j) = \sum_{\mu\nu} P_{\mu\nu}(\mathbf{B}, \mathbf{m}_j) \left[h_{\mu\nu}(\mathbf{B}, \mathbf{m}_j) + \frac{1}{2} G_{\mu\nu}(\mathbf{B}, \mathbf{m}_j) \right], \quad (3.126)$$

with

$$h_{\mu\nu}(\mathbf{B}, \mathbf{m}_j) = \langle \phi_\mu(\mathbf{B}) | \mathbf{p}_B^2 - \sum_A \frac{Z_A}{\mathbf{r}_A} | \phi_\nu(\mathbf{B}) \rangle \quad (3.127)$$

$$\begin{aligned} G_{\mu\nu}(\mathbf{B}, \mathbf{m}_j) &= \mathbf{\Pi}(\mathbf{B}, \mathbf{m}) \mathbf{P}(\mathbf{B}, \mathbf{m}) \\ &= (\phi_\mu(\mathbf{B}) \phi_\nu(\mathbf{B}) || \phi_\lambda(\mathbf{B}) \phi_\sigma(\mathbf{B})) P_{\lambda\sigma}(\mathbf{B}, \mathbf{m}_j). \end{aligned} \quad (3.128)$$

3.6.2.2 Explicit Expressions for the Different Terms in the Perturbative Expansion

In order to obtain explicit expressions for the terms in eq. 3.117 we have to calculate the derivatives of the corresponding matrix elements. Since the GIAOs introduce a B-field dependence into the basis functions, the one-electron integral derivatives can be split by the product rule into the matrix representation of the derivative of the operator \hat{O} and terms resulting from the derivation of the product of gauge factors

$$\begin{aligned} \frac{\partial}{\partial B_i} \langle \chi_\mu(\mathbf{B}) | \hat{O} | \chi_\nu(\mathbf{B}) \rangle &= \langle \frac{\partial}{\partial B_i} \chi_\mu(\mathbf{B}) | \hat{O} | \chi_\nu(\mathbf{B}) \rangle + \langle \chi_\mu(\mathbf{B}) | \frac{\partial}{\partial B_i} \hat{O} | \chi_\nu(\mathbf{B}) \rangle \\ &+ \langle \chi_\mu(\mathbf{B}) | \hat{O} | \frac{\partial}{\partial B_i} \chi_\nu(\mathbf{B}) \rangle \\ &= \langle \chi_\mu(0) | \frac{\partial}{\partial B_i} \hat{O} | \chi_\nu(0) \rangle + \frac{1}{2} \langle (\mathbf{R}_{\mu\nu} \times \mathbf{r}'_i) \chi_\mu(0) | \hat{O} | \chi_\nu(0) \rangle \\ &+ \frac{1}{2} (\mathbf{R}_\mu \times \mathbf{R}_\nu) \langle \chi_\mu(0) | \hat{O} | \chi_\nu(0) \rangle, \end{aligned} \quad (3.129)$$

with $\mathbf{R}_{\mu\nu} = \mathbf{R}_\mu - \mathbf{R}_\nu$. The variable \mathbf{r}' has a similar meaning as in the formulation of one-electron expectation values in terms of the 1-particle reduced density matrix in a

continuous basis, i.e. the operator \hat{O} does not act on \mathbf{r}' which will be substituted by \mathbf{r} before integration. Developing the Hamiltonian in eq. 3.122 in a Taylor series yields

$$\begin{aligned} \hat{H}_{\alpha\beta} = \hat{H}^{(0)} + \hat{h}_{\alpha}^{(1,0)} \Big|_{B_{\alpha}=0} B_{\alpha} + \sum_j \hat{h}_{\beta,j}^{(0,1)} \Big|_{m_{\beta,j}=0} m_{\beta,j} \\ + \sum_j \hat{h}_{\alpha\beta,j}^{(1,1)} \Big|_{B_{\alpha}, m_{\beta,j}=0} B_{\alpha} m_{\beta,j} + \dots, \end{aligned} \quad (3.130)$$

with $\alpha, \beta = x, y, z$ and the following one-electron operators

$$\hat{h}_{\alpha}^{(1,0)} = \frac{\partial \hat{h}}{\partial B_{\alpha}} = -\frac{i}{2} [(\mathbf{r} - \mathbf{R}_0) \times \mathbf{p}]_{\alpha}, \quad (3.131)$$

$$\hat{h}_{\beta,j}^{(0,1)} = \frac{\partial \hat{h}}{\partial m_{\beta,j}} = -i\alpha^2 \frac{[(\mathbf{r} - \mathbf{R}_j) \times \mathbf{p}]_{\beta}}{|\mathbf{r} - \mathbf{R}_j|^3}, \quad (3.132)$$

$$\hat{h}_{\alpha\beta,j}^{(1,1)} = \frac{\partial^2 \hat{h}}{\partial B_{\alpha} \partial m_{\beta,j}} = \frac{\alpha^2 (\mathbf{r} - \mathbf{R}_0)(\mathbf{r} - \mathbf{R}_j) \delta_{\alpha\beta} - (\mathbf{r} - \mathbf{R}_0)_{\alpha} (\mathbf{r} - \mathbf{R}_j)_{\beta}}{2 |\mathbf{r} - \mathbf{R}_j|^3}, \quad (3.133)$$

which we need to form the derivatives^{e)} in eq. 3.117 and the RHS of the CPSCF-equations (see section 3.6.1). Eq. 3.131 and 3.132 are the paramagnetic interaction operators which couple the external B-field (orbital angular momentum operator) and the nuclear magnetic moments (spin-orbit interaction operator) to the motion of the electrons, respectively. The mixed derivative second order term 3.133 is the diamagnetic interaction operator. For a general discussion of the operators and matrix elements given above and their physical interpretation see e.g. Ref. [102].

Inserting these expansions into Eq. 3.126 and ordering with respect to \mathbf{B} and \mathbf{m}_j yields the equations for determining the different derivatives of $E(\mathbf{B}, \mathbf{m}_j)$. Thus differentiating eq. 3.126 with respect to $m_{\beta,j}$ gives

$$\begin{aligned} E_{\beta,j}^{(0,1)} &= \sum_{\mu\nu} P_{\mu\nu,\beta,j}^{(0,1)} \left[h_{\mu\nu}^{(0)} + \frac{1}{2} G_{\mu\nu}^{(0)} \right] + P_{\mu\nu}^{(0)} \left[h_{\mu\nu,\beta,j}^{(0,1)} + \frac{1}{2} G_{\mu\nu,\beta,j}^{(0,1)} \right] \\ &= \sum_{\mu\nu} P_{\mu\nu}^{(0)} h_{\mu\nu,\beta,j}^{(0,1)}, \end{aligned} \quad (3.134)$$

where we used the derivative of the orthogonality relation and the fact, that the basis functions do not depend on the nuclear magnetic moments. The derivative with respect to the magnetic nuclear moments leads to

$$\frac{\partial h_{\mu\nu}}{\partial m_{\beta,j}} = h_{\mu\nu,\beta,j}^{(0,1)} = -i\alpha^2 \langle \chi_{\mu}(0) | \frac{[\mathbf{r}_j \times \nabla]_{\beta}}{r_j^3} | \chi_{\nu}(0) \rangle, \quad (3.135)$$

^{e)}The shorthand notation for the derivatives in Eqs. 3.130-3.133 will be used throughout this section.

with $\mathbf{r}_j = \mathbf{r} - \mathbf{R}_j$.

The second order term in eq. 3.117 is the sum of the expectation value of operator 3.133 and the derivative of Eq. 3.135 with respect to the external magnetic field \mathbf{B} :

$$\frac{d}{dB_\alpha} h_{\mu\nu,\beta,j}^{(0,1)} = \frac{\alpha^2}{2} \langle (\mathbf{R}_{\mu\nu} \times \mathbf{r}')_\alpha \chi_\mu(0) | \frac{[\mathbf{r}_j \times \nabla]_\beta}{r_j^3} | \chi_\nu(0) \rangle. \quad (3.136)$$

We obtain:

$$\begin{aligned} h_{\mu\nu,\alpha\beta,j}^{(1,1)} &= \frac{\alpha^2}{2} \langle \chi_\mu(0) | \frac{(\mathbf{r} - \mathbf{R}_0)(\mathbf{r} - \mathbf{R}_j)\delta_{\alpha\beta} - (\mathbf{r} - \mathbf{R}_0)_\alpha(\mathbf{r} - \mathbf{R}_j)_\beta}{r_j^3} | \chi_\nu(0) \rangle \\ &+ \frac{\alpha^2}{2} \langle (\mathbf{R}_{\mu\nu} \times \mathbf{r}')_\alpha \chi_\mu(0) | \frac{[\mathbf{r}_j \times \nabla]_\beta}{r_j^3} | \chi_\nu(0) \rangle. \end{aligned} \quad (3.137)$$

To build the RHS of the CPSCF-equations we also have to know the expressions for the magnetic field derivatives, where we have to consider the explicit field dependence of the basis functions:

$$\begin{aligned} h_{\mu\nu,\alpha}^{(1,0)} &= -\frac{i}{2} \left[\langle \chi_\mu(0) | [\mathbf{r} \times \nabla]_\alpha | \chi_\nu(0) \rangle + \langle (\mathbf{R}_{\mu\nu} \times \mathbf{r}')_\alpha \chi_\mu(0) | \nabla^2 | \chi_\nu(0) \rangle \right. \\ &\left. + \sum_A \langle (\mathbf{R}_{\mu\nu} \times \mathbf{r}')_\alpha \chi_\mu(0) | Z_A r_A^{-1} | \chi_\nu(0) \rangle \right]. \end{aligned} \quad (3.138)$$

The derivatives of the 2-electron integrals are

$$\begin{aligned} G_{\mu\nu,\alpha}^{(1,0)} &= \sum_{\lambda\sigma} P_{\lambda\sigma,\alpha}^{(1,0)} \left[(\chi_\mu(0)\chi_\nu(0) || \chi_\lambda(0)\chi_\sigma(0)) \right] \\ &+ \frac{i}{2} \sum_{\lambda\sigma} P_{\lambda\sigma} \left[((\mathbf{R}_{\mu\nu} \times \mathbf{r}'_1)_\alpha \chi_\mu(0)\chi_\nu(0) || (\mathbf{R}_{\lambda\sigma} \times \mathbf{r}'_2)_\alpha \chi_\lambda(0)\chi_\sigma(0)) \right]. \end{aligned} \quad (3.139)$$

3.6.2.3 GIAO Kohn-Sham DFT

For GIAO-KS-DFT we will only discuss the well-established LSDA, GGA and Hybrid XC functionals and exclude current-dependent functionals. Using KS-DFT we have to consider the perturbation on $\mathbf{G}_{DFT}(\mathbf{P})$ because of the B-field dependence of the basis functions. Since the non-hybrid XC functionals are local in the density (i.e. in contrast to the non-local HF exchange) the terms $\mathbf{G}_{DFT}(\mathbf{P}^x)$ vanish while hybrid functionals still contain some "exact" exchange. Thus the only contribution to the two-electron terms is due to the GIAOs in $\mathbf{G}_{DFT}^x(\mathbf{P})$, which is for LSDA functionals

$$\frac{\partial}{\partial B_x} \langle \chi_\mu(\mathbf{B}) | \frac{\partial E_{xc}}{\partial \rho} | \chi_\nu(\mathbf{B}) \rangle = \frac{i}{2} \int \frac{\partial E_{xc}}{\partial \rho(\mathbf{r})} (\mathbf{R}_{AB} \times \mathbf{r})_x [\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})] d\mathbf{r}. \quad (3.140)$$

For GGA functionals we obtain

$$\begin{aligned}
 \frac{\partial}{\partial B_x} \langle \chi_\mu(\mathbf{B}) | \frac{\partial E_{xc}}{\partial \rho} | \chi_\nu(\mathbf{B}) \rangle &= \frac{i}{2} \int \frac{\partial E_{xc}}{\partial \rho(\mathbf{r})} (\mathbf{R}_{AB} \times \mathbf{r})_x [\chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r})] d\mathbf{r} \\
 &+ \frac{i}{2} \int \frac{\partial E_{xc}}{\partial |\nabla \rho(\mathbf{r})|} \frac{\nabla \rho}{|\nabla \rho|} (\mathbf{R}_{AB} \times \mathbf{r})_x [\nabla \chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r})] d\mathbf{r} \\
 &+ \frac{i}{2} \int \frac{\partial E_{xc}}{\partial |\nabla \rho(\mathbf{r})|} \left(\mathbf{R}_{AB} \times \frac{\nabla \rho}{|\nabla \rho|} \right)_x [\chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r})] d\mathbf{r}.
 \end{aligned} \tag{3.141}$$

As it is easily seen, the B-field derivatives only need slight modifications in the numerical integration routines of an existing KS-DFT code.

3.6.3 Implementational Details

Here only the implementation of perturbed integrals and the linear equation solver are briefly discussed, for informations on the sparse algebra routines see the following chapter.

3.6.3.1 Linear Equation Solver

We solve eq. 3.98 and eq. 3.105 within two steps (Level-1 and -2) by splitting the LHS of our linear equation system into two parts \mathbf{A}_1 and \mathbf{A}_2 which is shown for eq. 3.98:

$$\mathbf{A}\mathbf{x} = (\mathbf{A}_1 + \mathbf{A}_2)\mathbf{x} = \mathbf{b}, \tag{3.142}$$

with

$$\begin{aligned}
 \mathbf{A}_1\mathbf{x} &= \mathbf{F}\mathbf{P}^x\mathbf{S}\mathbf{P}\mathbf{S} - \mathbf{F}\mathbf{P}\mathbf{S}\mathbf{P}^x\mathbf{S} + \mathbf{S}\mathbf{P}\mathbf{S}\mathbf{P}^x\mathbf{F} - \mathbf{S}\mathbf{P}^x\mathbf{S}\mathbf{P}\mathbf{F}, \\
 \mathbf{A}_2\mathbf{x} &= \mathbf{G}[\mathbf{P}^x]\mathbf{P}\mathbf{S} + \mathbf{S}\mathbf{P}\mathbf{G}[\mathbf{P}^x] - 2\mathbf{S}\mathbf{P}\mathbf{G}[\mathbf{P}^x]\mathbf{P}\mathbf{S}.
 \end{aligned} \tag{3.143}$$

At Level-1 we solve the equation

$$\begin{aligned}
 \mathbf{A}_1\mathbf{x} &= \mathbf{b} - \mathbf{A}_2\mathbf{x}, \\
 \implies \mathbf{A}_1\mathbf{x} &= \tilde{\mathbf{b}},
 \end{aligned} \tag{3.144}$$

with a conjugate gradient routine for a positive definite matrix \mathbf{A} . See also the scheme in Fig. B.5. The extension to a biconjugate gradient routine for non-positive definite matrices improves the convergence behavior only slightly, but it also requires a second formation of $\mathbf{A}_1\mathbf{x}$ and so doubles the number of the dominating multiplications in each iteration.

Therefore, the simple conjugate gradient algorithm was chosen as standard method in our current implementation.

At Level-2 we solve the complete equation where the convergence is accelerated by a modified Pulay's DIIS method [104] using the residual $\mathbf{b} - \mathbf{A}\mathbf{x}$ as error vector. Apart from the extrapolation of \mathbf{P}^x and $\mathbf{A}_2\mathbf{x}$, the corresponding residual $\mathbf{b} - \mathbf{A}_1\mathbf{x}$ is also extrapolated. Thus the computation of the LHS with the complete perturbed density matrix \mathbf{P}^x is avoided in the determination of the error $\mathbf{b} - \mathbf{A}\mathbf{x}$.

Using the contravariant LEQS (eq. 3.105) we can avoid the transformation in eq. 3.103 at each Level-1 step. Instead, we have to consider multiplications with the sparse metric \mathbf{S} in order to determine the step length of the conjugate gradient step (see e.g. Ref. [105]). Furthermore, a reprojection onto the occ/virt and virt/occ subspaces is unnecessary since it is implied in the structure of our LEQS.

3.6.3.2 Integral Engines

All integral routines are based on an Obara-Saika scheme [106] combined – if possible – with the horizontal recursion of Head-Gordon and Pople [107].

The angular momentum integrals in eq. 3.131 are computed as a linear combination of overlap integrals, the derivatives of the kinetic, nuclear-attraction in eq. 3.138 and electron-electron repulsion integrals in eq. 3.139 only require an additional computation of the expectation values with the angular momentum quantum number raised by 1. The electric field integrals necessary for forming the matrix of operator in eq. 3.133 and the spin-orbit interaction integrals as well as their B-field derivatives can also be easily implemented. Since the corresponding operators can be formulated as derivatives of the Coulomb operator, one only has to modify the already existing nuclear-attraction routines.

To achieve a linear scaling behavior we need to adapt CFMM [1] and LinK [8] to the perturbed Coulomb and exchange matrix, respectively. As mentioned before, the CFMM scheme is based on a separation of the interacting electron field into a near-field (NF) and a far-field part (FF). While the NF-interaction is treated within a standard integral engine the FF is described by a multipole field. If a standard routine to calculate the unperturbed Coulomb matrix is implemented, only the far field has to be modified. From the skew-symmetry of the prefactor matrix in eq. 3.139 it is obvious, that the terms in the summation over the perturbed ket-side cancel each other, so we only have to consider the

perturbation on the bra-side:

$$\begin{aligned} J_{\mu\nu}^x(\mathbf{P}) &= \sum_{\lambda\sigma} P_{\lambda\sigma} \left[\left((\mathbf{R}_{\mu\nu} \times \mathbf{r}'_1)_x \chi_\mu(0)\chi_\nu(0) \left| r_{12}^{-1} \right| (\mathbf{R}_{\lambda\sigma} \times \mathbf{r}'_2)_x \chi_\lambda(0)\chi_\sigma(0) \right) \right] \\ &= \sum_{\lambda\sigma} P_{\lambda\sigma} \left[\left((\mathbf{R}_{\mu\nu} \times \mathbf{r}'_1)_x \chi_\mu(0)\chi_\nu(0) \left| r_{12}^{-1} \right| \chi_\lambda(0)\chi_\sigma(0) \right) \right]. \end{aligned} \quad (3.145)$$

For similar reasons, i.e. the skew-symmetry of the perturbed density \mathbf{P}^x , the Coulomb-type contractions with the two-electron integrals vanish:

$$J_{\mu\nu}(\mathbf{P}^x) = \sum_{\lambda\sigma} P_{\lambda\sigma}^x (\chi_\mu(0)\chi_\nu(0) | \chi_\lambda(0)\chi_\sigma(0)) = 0. \quad (3.146)$$

The perturbed exchange matrices $\mathbf{K}^x(\mathbf{P})$ scale naturally linear with system size for a system with a non-vanishing band gap because of the coupling of the electron distributions $\chi_\mu\chi_\nu$ and $\chi_\lambda\chi_\sigma$ by the elements of the density matrix \mathbf{P} . This is hampered by the quadratical scaling of traditional screening routines, a disadvantage that is overcome by the LinK method [8, 9]. Since the screening focuses mainly on the density matrix, a small number of modifications is needed to form the B-field derivatives of the exchange matrix:

$$\begin{aligned} K_{\mu\nu}^x(\mathbf{P}) &= \frac{i}{2} \sum_{\nu\sigma} P_{\nu\sigma} \\ &\quad \times \left((\mathbf{R}_{\mu\nu} \times \mathbf{r}'_1)_x \chi_\mu\chi_\nu \left| (\mathbf{R}_{\lambda\sigma} \times \mathbf{r}'_2)_x \chi_\lambda\chi_\sigma \right. \right), \end{aligned} \quad (3.147)$$

$$K_{\mu\nu}(\mathbf{P}^x) = \sum_{\nu\sigma} P_{\nu\sigma}^x (\chi_\mu\chi_\nu | \chi_\lambda\chi_\sigma). \quad (3.148)$$

While the exchange-like contraction of the ground-state density \mathbf{P} with the perturbed integrals within a LinK algorithm is straightforward (eq. 3.147), the exchange matrices built from the perturbed densities \mathbf{P}^x are required for each Level-2 cycle (eq. 3.148). Here at most three matrices have to be constructed, so that – instead of creating only one minilist in the screening process – we form three lists corresponding to the different perturbed densities and merge them to yield a complete list of significant shell pairs to be calculated. In the final step we again use the single minilists to ensure a linear scaling behavior for the contraction with the different \mathbf{P}^x matrices.

3.6.4 Applications of D-GIAO-CPSCF

As a first example, the scaling behavior of our new D-GIAO-HF method is shown for a series of linear alkanes using a 6-31G* basis (Fig. B.3). For all computations the same

integral thresholds ($> 10^{-6}$) and convergence criteria ($\|RHS - LHS\| < 10^{-3}$) are chosen which provide numerical accuracies better than 0.1 ppm for proton chemical shifts.

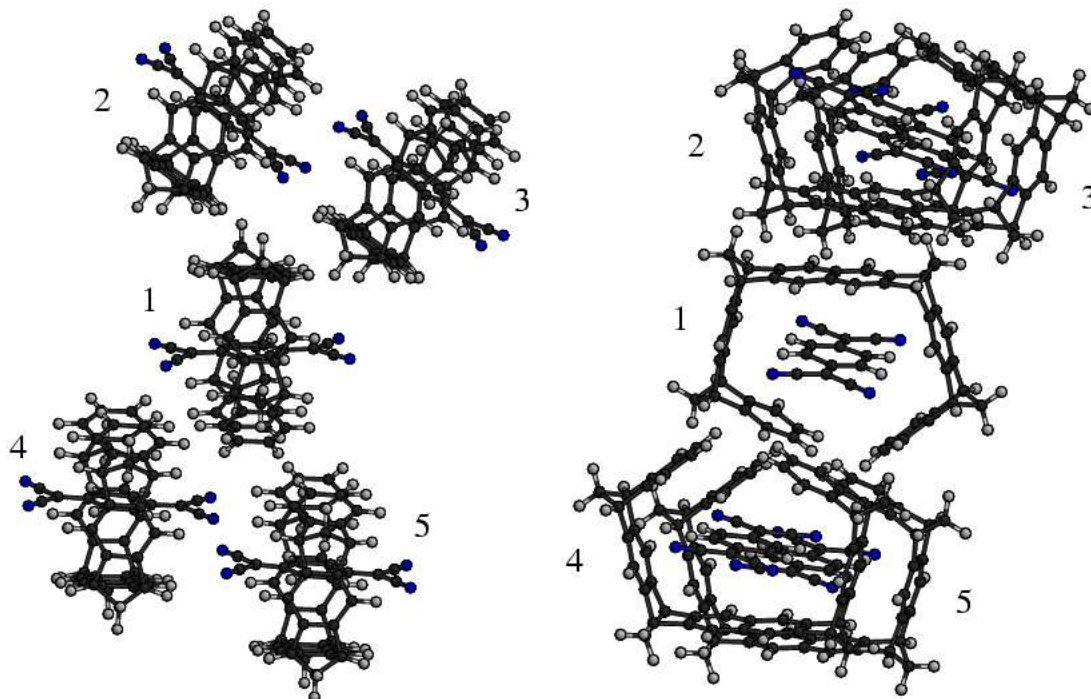
In Fig. B.3 in the appendix one can see that the application of the CFMM and LinK (triangles) clearly improves the performance compared to the standard $\mathcal{O}(M^3)$ algorithm (cubes) by a factor of approx. 2.3 for $C_{160}H_{322}$. The linear scaling behavior of the integral routines hides the cubic scaling of the algebra routines until approx. 1150 basis functions, then the $\mathcal{O}(M^3)$ matrix multiplications start to become significant. If sparse matrix algebra routines (circles) are applied an overall $\mathcal{O}(M)$ scaling behavior is reached and an overall enhancement factor of approx. 3.9 for $C_{160}H_{322}$ compared to the standard routines.

The performance of the D-GIAO-KS-DFT algorithm is shown in Fig. B.4 in the appendix for the computation of a series of amylose chains. It has to be mentioned that for XC functionals without exact exchange, the MO-based schemes only require a single construction of \mathbf{U}^x according to eq. 3.64 since $\mathbf{G}(\mathbf{P}^x)$ vanishes. Using the present D-CPSCF algorithm (eq. 3.98) in contrast, we have to solve the LEQS at Level-1 once ($\mathbf{A}_1\mathbf{x} = \mathbf{b} - \mathbf{A}_2\mathbf{x}$, see sec. 3.6.3.1). While using a tight convergence criterion, we obtain a similar picture as in the case of D-GIAO-HF for linear alkanes, i.e. an overall linear scaling is obtained by application of $\mathcal{O}(M)$ integral contractions in combination with sparse algebra routines.

With this favorable scaling behavior the calculation of system sizes in the 1000 atoms region becomes possible. As a proof of principle the application to solid- and solution-state problems is presented. It has been shown in several examples, e.g. Refs. [16–19] that the reliable assignment of experimental NMR spectra for solid-state systems has become possible by quantum chemical calculations. In such examples it is crucial to converge the theoretically determined NMR chemical shifts with the size of the solid-state fragment. However, because of the cubically increasing effort in computational time of standard GIAO-HF or GIAO-DFT methods, such calculations have been restricted to the 100 atom region. Therefore, one had to resort to small basis sets and incremental approaches for treating larger compounds.

With this method Ochsenfeld et al. [16] investigated a host-guest complex (Fig. 3.1) with a tweezer-shaped host that binds a tetracyano-*p*-quinodimethane (TCNQ). The ^1H -NMR chemical shifts for the central monomer in Fig. 3.1 have been estimated from trimer values (**1**, **2**, **3** and **1**, **4**, **5**, each containing 276 atoms). Therefore, the influences of the complexes **2**, **3** as well as **4**, **5** on the central complex **1** have been determined from trimer

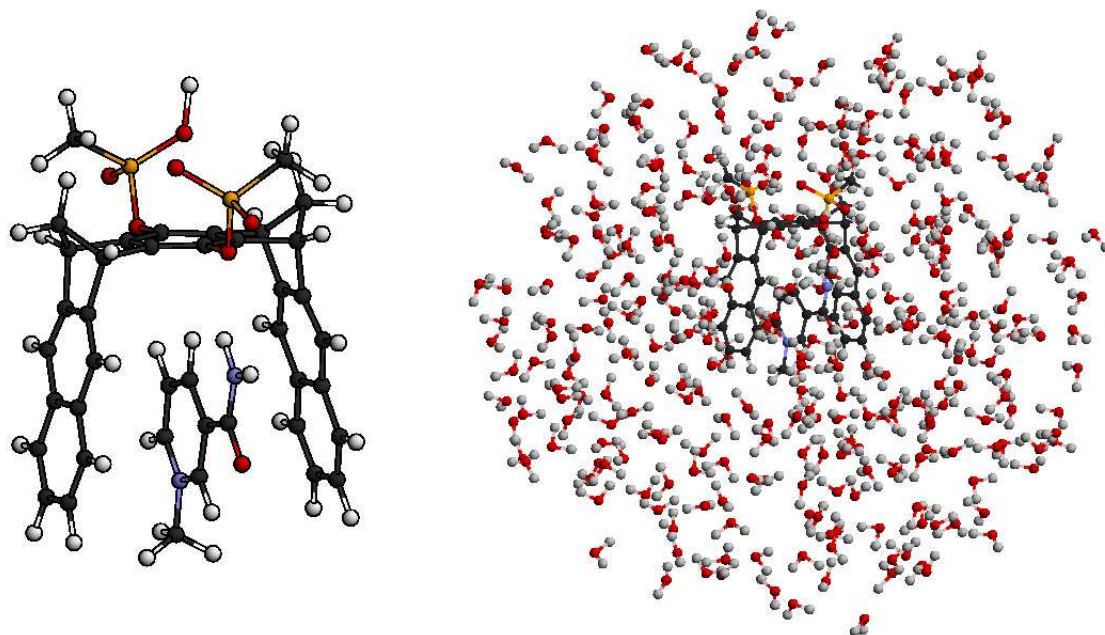
Figure 3.1: Pentamer of host-guest complex.



calculations with GIAO-HF/3-21G ($\Delta\delta(3-21G)$) and were then added to the corresponding monomer data ($\delta(3-21G)$). The error of this incremental approach compared to a pentamer calculation with D-GIAO-HF/3-21G was determined to be in the order of 0.3 ppm, showing that the incremental approach yields proper results. Nevertheless, the 3-21G basis is too small to provide reliable results. Comparing the incremental results ($\Delta\delta(3-21G) + \delta(3-21G)$) with those of a pentamer calculation using D-GIAO-HF/6-31G*, the error is in the order of 0.8 ppm. This error can be slightly lowered to 0.7 ppm by considering the influences on the monomer shifts in an incremental fashion starting from a monomer calculation with a 6-31G* basis and trimer calculations with a 3-21G basis. With this incremental approach the mean error is reduced from 0.5 to 0.2 ppm. Nevertheless, the calculation of larger fragments with reasonable basis sets, which becomes possible by the newly developed linear scaling NMR method, is certainly the most rigorous way and improves the accuracy considerably.

As a second example we converged the NMR chemical shifts of a host-guest complex in water with respect to the size of the surrounding hydration shell. The system of in-

Figure 3.2: Snapshot of molecular dynamics simulation of *N*-methyl nicotinamide in water.



terest is an artificial receptor for NAD⁺ (nicotinamide adenine dinucleotide) which has been presented by Klärner and Schrader [20]. The structure of this complex could not be revealed by experimental informations only, thus an additional investigation with theoretical methods has been employed in order to determine possible binding motifs. Since the experimental spectra have been measured in water, we will only focus on the effect of the hydration shell on the chemical shielding constants of the guest molecule. In order to elucidate the effect of the hydration shell, a molecular dynamics (MD) simulation (Force field: MMFF94, T = 300 K, 1 ps equilibration, total simulation time: 20 ps with $\Delta t = 1$ fs) of *N*-methyl nicotinamide bound to the host molecule (Fig. 3.2) has been performed. From the different configurations the snapshot has been selected, for which the water molecules are closest to the center of the ring of the guest molecule within a radius of 8 Å. From this single snapshot spherical water shells with different radii have been cut out. For the different fragments D-GIAO-HF/6-31G** calculations have been performed where the largest system contained 1003 atoms. As can be seen from Tab. 3.2 the NMR shifts of the atoms of the host-guest complex are converged to 0.2 ppm with respect of the radius of the hy-

Table 3.2: Max. change of NMR shifts of the host-guest system with increasing hydration shell.

Change in System Size (No. Atoms)	Max. Change [ppm]
88 → 169	1.3
169 → 547	1.1
547 → 1003	0.2

dration shell. It has to be stressed that this is only a proof of principle that calculations of large systems is possible with our new linear scaling method. For an accurate investigation of the solvent effects of course multiple snapshots of a MD simulation would need to be considered. However, this is not feasible on todays workstations and here we were mainly interested in estimating a maximum effect and the size of the required hydration shell.

3.7 Linear Scaling Methods for Second and Third Order Properties: First and Second Order Density Matrix-based TDSCF

In this section we will discuss the calculation of the linear and quadratic response to an oscillating external electric field. Since the corresponding properties are characterized by electronic transitions between different states, the derivative scheme for a definite state, as presented above for the calculation of NMR spectra, is not applicable.

In the following we will present a time-dependent theory defined in terms of the 1-particle reduced density, which is both straightforward and overcomes some of the obstacles of standard MO-based methods [84, 86].

3.7.1 Linear Response Equations for a Dynamic Perturbation

The first derivative of the time-dependence of the density in eq. 2.45 with respect to the electric field strength \mathcal{E}_x is given by

$$\begin{aligned} \mathbf{F}^x \mathbf{P} \mathbf{S} + \mathbf{F} \mathbf{P}^x \mathbf{S} - \mathbf{S} \mathbf{P}^x \mathbf{F} - \mathbf{S} \mathbf{P} \mathbf{F}^x &= i \dot{\mathbf{S}} \mathbf{P}^x \mathbf{S} \\ &= \kappa \omega \mathbf{S} \mathbf{P}^x \mathbf{S}. \end{aligned} \quad (3.149)$$

In the second line the time derivative of the exponential factor has been applied where κ equals zero in the static case, and +1 and -1 for the terms $e^{-i\omega t}$ and $e^{+i\omega t}$, respectively (see the last column in Tab. 3.1 (p. 40) for the corresponding values).

In order to obtain the D-TDSCF equations one has to impose the constraints for \mathbf{P}^x that are given in sec. 3.4. This is done in the same fashion as shown for the static response equations, i.e. by the difference between the half-projected equations of eq. 3.149:

$$\left(\text{LEQS}_{eq. 3.149} \right) \mathbf{P} \mathbf{S} - \mathbf{S} \mathbf{P} \left(\text{LEQS}_{eq. 3.149} \right). \quad (3.150)$$

With the explicit expressions

$$\mathbf{S} \mathbf{P} \mathbf{F}^x \mathbf{P} \mathbf{S} + \mathbf{S} \mathbf{P} \mathbf{F} \mathbf{P}^x \mathbf{S} - \mathbf{S} \mathbf{P} \mathbf{S} \mathbf{P}^x \mathbf{F} - \mathbf{S} \mathbf{P} \mathbf{F}^x \mathbf{S} = \kappa \omega \mathbf{S} \mathbf{P} \mathbf{S} \mathbf{P}^x \mathbf{S}, \quad (3.151)$$

$$\mathbf{F}^x \mathbf{P} \mathbf{S} + \mathbf{F} \mathbf{P}^x \mathbf{S} \mathbf{P} \mathbf{S} - \mathbf{S} \mathbf{P}^x \mathbf{F} \mathbf{P} \mathbf{S} - \mathbf{S} \mathbf{P} \mathbf{F}^x \mathbf{P} \mathbf{S} = \kappa \omega \mathbf{S} \mathbf{P}^x \mathbf{S} \mathbf{P} \mathbf{S}, \quad (3.152)$$

one can set up an effective LEQS for \mathbf{P}^x :

$$\begin{aligned}
 & \mathbf{FP}^x\mathbf{SPS} - \mathbf{FPSP}^x\mathbf{S} + \mathbf{SPSP}^x\mathbf{F} - \mathbf{SP}^x\mathbf{SPF} + \mathbf{G}(\mathbf{P}^x)\mathbf{PS} + \mathbf{SPG}(\mathbf{P}^x) \\
 & - 2\mathbf{SPG}(\mathbf{P}^x)\mathbf{PS} - \kappa\omega(\mathbf{SP}^x\mathbf{SPS} - \mathbf{SPSP}^x\mathbf{S}) \\
 = & -\mathbf{F}^{(x)}\mathbf{PS} - \mathbf{SPF}^{(x)} + 2\mathbf{SPF}^{(x)}\mathbf{PS},
 \end{aligned} \tag{3.153}$$

with the matrix representation of the dipole moment operator μ^x

$$F_{\mu\nu}^{(x)} = \mu_{\mu\nu}^x. \tag{3.154}$$

Note that we already sorted the LEQS by shifting all terms containing \mathbf{P}^x to the LHS. As in the static case, the constraints on \mathbf{P}^x are imposed only by the projection properties of the ground-state density. Thus one solves iteratively only for the occ/virt and virt/occ part of \mathbf{P}^x while the occ/occ part \mathbf{P}_{oo}^x is given as initial guess and the virt/virt part \mathbf{P}_{vv}^x vanishes. Furthermore a comparison with the LEQS for the calculation of static 2nd order properties in eq. 3.98 shows the equality for the zero-frequency limit (if $\mathbf{S}^x = \mathbf{0}$).

Note that the antisymmetry of \mathbf{P}^x as expected for ($\omega \neq 0$) is ensured by the last term $\kappa\omega(\mathbf{SP}^x\mathbf{SPS} - \mathbf{SPSP}^x\mathbf{S})$ on the LHS of eq. 3.153. If the initial transition density is given as the symmetric, contravariant RHS of eq. 3.153, the last term of the LHS would be skew-symmetric and therefore an antisymmetric LHS would result. In the limit ($\omega \rightarrow 0$) we have a symmetric \mathbf{P}^x as expected for the static case.

One advantage of the density-matrix approach is already evident: We do not have to consider special algorithms for the calculation of polarizabilities at ω close to the orbital differences $\Delta\epsilon_{ai} = \epsilon_a - \epsilon_i$ of the unperturbed Hamiltonian. In standard implementations \mathbf{P}^x or \mathbf{U}^x , respectively, are updated by the spectral formula in eq. 3.52 where the single transition dyadics are weighted by $1/[\omega - \Delta\epsilon_{ai}]$. Thus we had to transform the system to a non-canonical representation for $\omega \approx \Delta\epsilon$ in order to avoid singularities if the standard MO-based algorithm is used.

3.7.1.1 Linear Response of the Exchange-Correlation Functional

For TDDFT calculations we have to consider the response in the exchange-correlation (XC) potential

$$v_{xc}(\mathbf{r}, t) = \frac{\partial E_{xc}}{\partial \rho(\mathbf{r}, t)}. \tag{3.155}$$

In a perturbative expansion of the XC functional an explicit time-dependence of the XC kernel f_{xc} in the first order term arises^{f)}

$$\begin{aligned} v_{xc}^x(\mathbf{r}, t) &= \int f_{xc}(\mathbf{r}, \mathbf{r}', t) \rho^x(\mathbf{r}', t) d\mathbf{r}' \\ &= \int \frac{\partial^2 E_{xc}}{\partial \rho(\mathbf{r}) \partial \rho(\mathbf{r}')} \rho^x(\mathbf{r}', t) d\mathbf{r}'. \end{aligned} \quad (3.156)$$

Within the adiabatic approximation, which is used in this work, it is assumed that the time-dependent XC functional depends on the time-dependent density in the same way as the time-independent functional on the time-independent density does. This approximation works well for slow, i.e. adiabatic, processes (e.g. Ref. [108]). As a consequence, the time-dependency of f_{xc} is neglected

$$v_{xc}^x(\mathbf{r}, t) = \int f_{xc}(\mathbf{r}, \mathbf{r}', 0) \rho^x(\mathbf{r}', t) d\mathbf{r}'. \quad (3.157)$$

Thus the D-TDDFT equations are obtained by replacing the two-electron term $\mathbf{G}(\mathbf{P}^x)$ in eq. 3.153 by

$$\mathbf{G}(\mathbf{P}^x) \rightarrow \mathbf{G}_{DFT}(\mathbf{P}^x) = \mathbf{J}(\mathbf{P}^x) + \mathbf{V}_{xc}^{(2)}(\mathbf{P}^x), \quad (3.158)$$

where $\mathbf{V}_{xc}^{(2)}(\mathbf{P}^x)$ is the matrix representation of $v_{xc}^x(\mathbf{r}, t)$ in the given basis. So far we assumed the response to be causal, but it has to be noted that a contradiction between the symmetry of the second order derivative of E_{xc} and causality requirements arises [74]. It can be easily seen that the linear response is invariant to the interchange of the order of differentiation

$$\frac{\partial}{\partial \rho(\mathbf{r}', t')} \frac{\partial E_{xc}}{\partial \rho(\mathbf{r}, t)} = \frac{\partial}{\partial \rho(\mathbf{r}, t)} \frac{\partial E_{xc}}{\partial \rho(\mathbf{r}', t')}, \quad (3.159)$$

but causality requires that the response is zero for $t < t'$. This problem can be resolved by a formulation within the Keldysh formalism [74].

3.7.1.2 Molecular Polarizability – Illustrative Examples

The scaling behavior of the first order D-TDSCF methods is shown for the example of static polarizabilities of a series of linear alkanes at the TDDFT-BP86(VWN)/6-31G* [109–111] level of theory in Fig. B.6. For all computations the same integral thresholds ($> 10^{-7}$) and

^{f)}See appendix A for an explicit expression.

Table 3.3: Excitation energies and oscillator strengths for the first three excited singlet states from a TDDFT calculation with BP86(VWN)/6-31G* of HCF₃.

Frequency [a.u.]/[eV]	Oscillator Strength f
0.3664 / 9.9713	0.0100
0.5593 / 15.2181	0.3009
1.0833 / 29.4769	0.0654

convergence criteria ($\|RHS - LHS\| < 10^{-4}$) are chosen. Note that the general routines for the treatment of frequency-dependent polarizabilities are used, i.e. the symmetry of the transition densities $\mathbf{P}^x(\omega = 0)$ have not been exploited. Because of the tight convergence criterion, the number of iterations in the conjugate gradient algorithm is quite large, thus the matrix multiplications clearly dominate the overall computational effort. Thus, as it is shown in Fig. B.6, the application of the CFMM method to determine $\mathbf{J}(\mathbf{P}^x)$ (triangles) is only a moderate improvement compared to the $\mathcal{O}(M^2)$ integral routine (cubes). The further application of sparse multiplications (circles) clearly reduces the overall computational time and provides a linear scaling behavior.

As a second example the polarizability of the hydrogen fluoride molecule (bond length: 0.9254 Å, orientation: z-axis) for different frequencies up to 1.1 a.u. (29.93 eV) at the BP86(VWN)/6-31G* level is presented. The plot for the different frequencies is depicted in Fig. B.7. On the left of Fig. B.7 the data points for all calculated frequencies are shown. The red areas denote the poles of the polarization propagator, i.e. the region close to one of the excitation frequencies shown in Tab. 3.3. In order to be able to interpret the results obtained for near-singular frequencies ω , the determination of excitation frequencies ω^{ex} has to be considered. Within the random phase approximation (RPA) [43, 112] we have to solve an eigenvalue equation

$$\mathbf{\Omega}\mathbf{U}(+\omega_i^{ex}) = \omega_i^{ex}\mathbf{U}(+\omega_i^{ex}) \quad (3.160)$$

with the excitation energies ω_i^{ex} as eigenvalues, \mathbf{U}_i^{ex} as the corresponding transition coefficients vectors, and $\mathbf{\Omega}$ denotes the Hessian matrix. In order to describe forced oscillations induced by an external field we have to solve

$$(\mathbf{\Omega} - \omega\mathbf{1})\mathbf{U} = \mathbf{H}^{(S)} \quad (3.161)$$

Table 3.4: Relative oscillator strengths of the three lowest excitations of HCF₃ with BP86(VWN)/6-31G*.

Rel. oscillator strengths	Excitation	Polarizability
$f_{0.5593}/f_{0.3664}$	30.09	29.35
$f_{0.5593}/f_{1.0833}$	4.60	6.30

where $\mathbf{H}^{(S)}$ denotes the interaction between the molecular system and the external field. To solve this LEQS we have to find the inverse $(\mathbf{\Omega} - \omega\mathbf{1})^{-1}$ (response matrix $\mathbf{\Pi}$). In the spectral representation this inverse is given as [43]

$$\mathbf{\Pi} = (\mathbf{\Omega} - \omega\mathbf{1})^{-1} = \sum_{\omega_i^{ex} > 0} \frac{\mathbf{U}(+\omega_i^{ex})\mathbf{U}^\dagger(+\omega_i^{ex})}{\omega_i^{ex} - \omega} + \frac{\mathbf{U}(-\omega_i^{ex})\mathbf{U}^\dagger(-\omega_i^{ex})}{\omega_i^{ex} + \omega}. \quad (3.162)$$

Thus the natural frequencies ω_i^{ex} are given by the poles of the response matrix in eq. 3.162, i.e. the amplitude of the forced oscillation diverges if ω coincides with ω_i^{ex} . Note that small differences in the Hartree-Fock solution of the different program packages result in large deviations in the polarizability tensor for $\omega \approx \omega_i^{ex}$. The α_{xx} ($= \alpha_{yy}$) component of the polarizability tensor of hydrogen fluoride calculated at the HF/6-31G* level for the lowest singlet transition frequency ($\omega = 0.4388$ a.u.), for example, deviates strongly for different ab initio packages:

Q-Chem (This work)	-32667.84 a.u.
TURBOMOLE [113]	-1699.62 a.u.
DALTON [114]	-34439.71 a.u.

Note that those deviations do not represent an error in itself since the results at the poles of the polarization propagator have to diverge. The results for smaller frequencies, i.e. far from the singularity, are of course the same for all program packages. Apart from the inaccuracies for ω close to an excitation frequency, the relations between the oscillator strengths f_{freq} of the different excitations in Tab. 3.3 are crudely reproduced. Even if these calculations for $\omega \approx \omega_i^{ex}$ provide less accurate results for the oscillator strength ratios compared to RPA calculations, they show the stability of our linear equation solver even for near-singular systems ($\det |\mathbf{\Omega} - \omega\mathbf{1}| \rightarrow 0$).

Table 3.5: Dynamic polarizabilities of PNA obtained with Q-Chem (this work) and TURBOMOLE with TDHF/6-31G* and TDDFT-B3LYP/6-31G* in 10^{-25} esu.

Method	Frequency [a.u.]/[eV]	This work	TURBOMOLE
TDHF	0.000/0.00	117.87	117.86
	0.024/0.65	118.41	118.41
	0.048/1.30	120.11	120.11
TDDFT	0.000/0.00	127.78	127.79
	0.024/0.65	128.69	128.69
	0.048/1.30	131.64	131.65

As mentioned before, one has to transform the unperturbed molecular orbitals to a non-canonical representation in the MO-based scheme if the frequency ω coincides with an orbital difference ($\epsilon_a - \epsilon_i$). This is not necessary for the present D-TDSCF method since the transition densities are not obtained via a spectral formula like eq. 3.52. The corresponding frequencies are colored red in the left picture of Fig. B.7.

In section 3.7.2.4 the results for the first order hyperpolarizabilities of para-nitroaniline (PNA) are presented. These are calculated by Wigner’s $(2n + 1)$ rule based on first results for $+\omega$ and -2ω . The frequency-dependent polarizabilities obtained with B3LYP/6-31G* [115, 116] and HF/6-31G* are given in Tab. C.2 and Tab. C.3. In order to check the correctness of these results they are compared to the polarizabilities obtained with TURBOMOLE [113, 117] in Tab. 3.5.

3.7.2 Quadratic Response Equations for a Dynamic Perturbation

The second derivative of the commutator in eq. 3.149 is

$$\begin{aligned}
 & \mathbf{F}^{xy}\mathbf{P}\mathbf{S} + \mathbf{F}^x\mathbf{P}^y\mathbf{S} + \mathbf{F}^y\mathbf{P}^x\mathbf{S} + \mathbf{F}\mathbf{P}^{xy}\mathbf{S} \\
 & - \mathbf{S}\mathbf{P}^{xy}\mathbf{F} - \mathbf{S}\mathbf{P}^x\mathbf{F}^y - \mathbf{S}\mathbf{P}^y\mathbf{F}^x - \mathbf{S}\mathbf{P}\mathbf{F}^{xy} \\
 & = i\dot{\mathbf{S}}\mathbf{P}^{xy}\mathbf{S} = \kappa\omega\mathbf{S}\mathbf{P}^{xy}\mathbf{S}.
 \end{aligned} \tag{3.163}$$

Following the same scheme as used for the linear response (eqs. 3.151-3.152) we obtain the projections

$$\begin{aligned} & \mathbf{SPF}^{xy}\mathbf{PS} + \mathbf{SPF}^x\mathbf{P}^y\mathbf{S} + \mathbf{SPF}^y\mathbf{P}^x\mathbf{S} + \mathbf{FPSP}^{xy}\mathbf{S} \\ & - \mathbf{SPSP}^{xy}\mathbf{F} - \mathbf{SPSP}^x\mathbf{F}^y - \mathbf{SPSP}^y\mathbf{F}^x - \mathbf{SPF}^{xy} \\ & = \kappa\omega\mathbf{SPSP}^{xy}\mathbf{S}, \end{aligned} \quad (3.164)$$

$$\begin{aligned} & \mathbf{F}^{xy}\mathbf{PS} + \mathbf{F}^x\mathbf{P}^y\mathbf{SPS} + \mathbf{F}^y\mathbf{P}^x\mathbf{SPS} + \mathbf{FP}^{xy}\mathbf{SPS} \\ & - \mathbf{SP}^{xy}\mathbf{SPF} - \mathbf{SP}^x\mathbf{F}^y\mathbf{PS} - \mathbf{SP}^y\mathbf{F}^x\mathbf{PS} - \mathbf{SPF}^{xy}\mathbf{PS} \\ & = \kappa\omega\mathbf{SP}^{xy}\mathbf{SPS}. \end{aligned} \quad (3.165)$$

The second order D-TDSCF equations are obtained by the difference (eq. 3.165-eq. 3.164)

$$\begin{aligned} & \mathbf{FP}^{xy}\mathbf{SPS} - \mathbf{FPSP}^{xy}\mathbf{S} + \mathbf{SPSP}^{xy}\mathbf{F} - \mathbf{SP}^{xy}\mathbf{SPF} \\ & + \mathbf{G}(\mathbf{P}^{xy})\mathbf{PS} + \mathbf{SPG}(\mathbf{P}^{xy}) - 2\mathbf{SPG}(\mathbf{P}^{xy})\mathbf{PS} \\ & - \kappa\omega[\mathbf{SP}^{xy}\mathbf{SPS} - \mathbf{SPSP}^{xy}\mathbf{S}] \\ & = \mathbf{SP}^x\mathbf{F}^y\mathbf{PS} + \mathbf{SP}^y\mathbf{F}^x\mathbf{PS} - \mathbf{SPSP}^x\mathbf{F}^y - \mathbf{SPSP}^y\mathbf{F}^x \\ & - \mathbf{F}^x\mathbf{P}^y\mathbf{SPS} - \mathbf{F}^y\mathbf{P}^x\mathbf{SPS} + \mathbf{SPF}^x\mathbf{P}^y\mathbf{S} + \mathbf{SPF}^y\mathbf{P}^x\mathbf{S}, \end{aligned} \quad (3.166)$$

where we used $\mathbf{G}(\mathbf{P}^{xy}) = \mathbf{F}^{xy}$ (dipole approximation). All terms containing \mathbf{P}^{xy} were transferred to the LHS. As has been shown in sec. 3.4, \mathbf{P}_{oo}^{xy} and \mathbf{P}_{vv}^{xy} are completely determined by the products of the first order transition densities, so these parts must be eliminated in the Level-1 part of the LHS which can be easily proven by setting $\mathbf{P}^{xy} = \mathbf{P}^x\mathbf{SP}^y + \mathbf{P}^y\mathbf{SP}^x$. Analogous to first order, eq. 3.166 reassembles the structure of $\mathbf{P}^{xy}\mathbf{SP} + \mathbf{PSP}^{xy}$ (eq. 3.93).

A comparison between the linear and quadratic TDSCF equations shows that the same routines for their solution can be used. Except for the RHS, only slight modifications of the two-electron term $\mathbf{G}(\mathbf{P}^{xy})$ in case of TDDFT (see sec. 3.7.2.2) are required.

3.7.2.1 Initial Guess for the Second Order Transition Density

In the case of the second order density matrix it is necessary to mention the form of its initial value $\mathbf{P}_{\text{initial}}^{xy}$ explicitly. From the second derivative of the idempotency relation in eq. 3.30, one could be encouraged to use $\mathbf{P}^x\mathbf{SP}^y + \mathbf{P}^y\mathbf{SP}^x$, but this would not work with

the presented LEQS. To find a suitable $\mathbf{P}_{\text{initial}}^{xy}$ one can reconsider the resolution of \mathbf{P}^{xy} into its subspace projections:

$$\mathbf{P}^{xy} = \mathbf{P}_{\text{oo}}^{xy} + \mathbf{P}_{\text{ov}}^{xy} + \mathbf{P}_{\text{vo}}^{xy} + \mathbf{P}_{\text{vv}}^{xy}. \quad (3.167)$$

Note that we solve iteratively for $\mathbf{P}_{\text{ov}}^{xy} + \mathbf{P}_{\text{vo}}^{xy}$ within a LEQS that is constructed as:

$$\begin{aligned} \text{LEQS} &= (\text{LEQS}_{\text{orig}}) \mathbf{P} \mathbf{S} - \mathbf{S} \mathbf{P} (\text{LEQS}_{\text{orig}}) \\ &= (\text{LEQS}_{\text{orig}})_{\text{vo}} - (\text{LEQS}_{\text{orig}})_{\text{ov}}, \end{aligned} \quad (3.168)$$

so that the intrinsic occ/occ parts of $\mathbf{P}_{\text{ov}}^{xy}$ and $\mathbf{P}_{\text{vo}}^{xy}$ cancel each other

$$\begin{aligned} \mathbf{P}_{\text{ov}}^{xy} - \mathbf{P}_{\text{vo}}^{xy} &= \mathbf{P} \mathbf{S} \mathbf{P}^{xy} - \mathbf{P} \mathbf{S} \mathbf{P}^{xy} \mathbf{S} \mathbf{P} - \mathbf{P}^{xy} \mathbf{S} \mathbf{P} + \mathbf{P} \mathbf{S} \mathbf{P}^{xy} \mathbf{S} \mathbf{P} \\ &= \mathbf{P} \mathbf{S} \mathbf{P}^{xy} - \mathbf{P}^{xy} \mathbf{S} \mathbf{P}, \end{aligned} \quad (3.169)$$

and thus we actually obtain $\mathbf{P}^{xy} \mathbf{S} \mathbf{P} + \mathbf{P} \mathbf{S} \mathbf{P}^{xy}$. From

$$\mathbf{P}_{\text{ov}}^{xy} = \mathbf{P}^{xy} \mathbf{S} \mathbf{P} - \mathbf{P} \mathbf{S} \mathbf{P}^{xy} \mathbf{S} \mathbf{P} = \mathbf{P}^{xy} \mathbf{S} \mathbf{P} - \mathbf{P}_{\text{oo}}^{xy}, \quad (3.170)$$

$$\mathbf{P}_{\text{vo}}^{xy} = \mathbf{P} \mathbf{S} \mathbf{P}^{xy} - \mathbf{P} \mathbf{S} \mathbf{P}^{xy} \mathbf{S} \mathbf{P} = \mathbf{P} \mathbf{S} \mathbf{P}^{xy} - \mathbf{P}_{\text{oo}}^{xy}, \quad (3.171)$$

$$\implies \mathbf{P}^{xy} = \mathbf{P}_{\text{vv}}^{xy} + \mathbf{P} \mathbf{S} \mathbf{P}^{xy} + \mathbf{P}^{xy} \mathbf{S} \mathbf{P} - \mathbf{P}_{\text{oo}}^{xy}, \quad (3.172)$$

one can see that we actually require $-\mathbf{P}_{\text{oo}}^{xy}$ instead of $\mathbf{P}_{\text{oo}}^{xy}$. Thus we find as initial value

$$\mathbf{P}_{\text{initial}}^{xy} = (\mathbf{P}^x \mathbf{S} \mathbf{P}^y + \mathbf{P}^y \mathbf{S} \mathbf{P}^x)_{\text{vv}} - (\mathbf{P}^x \mathbf{S} \mathbf{P}^y + \mathbf{P}^y \mathbf{S} \mathbf{P}^x)_{\text{oo}} \quad (3.173)$$

$$\begin{aligned} &= (\mathbf{P}^x \mathbf{S} \mathbf{P} \mathbf{S} \mathbf{P}^y + \mathbf{P}^y \mathbf{S} \mathbf{P} \mathbf{S} \mathbf{P}^x) - (\mathbf{P} \mathbf{S} \mathbf{P}^x \mathbf{S} \mathbf{P}^y \mathbf{S} \mathbf{P} + \mathbf{P} \mathbf{S} \mathbf{P}^y \mathbf{S} \mathbf{P}^x \mathbf{S} \mathbf{P}) \\ &= \mathbf{P}^x \mathbf{S} \mathbf{P}^y + \mathbf{P}^y \mathbf{S} \mathbf{P}^x - 2 \mathbf{P} \mathbf{S} (\mathbf{P}^x \mathbf{S} \mathbf{P}^y + \mathbf{P}^y \mathbf{S} \mathbf{P}^x). \end{aligned} \quad (3.174)$$

3.7.2.2 Quadratic Response of the Exchange-Correlation Functional

A further differentiation of eq. 3.156 yields the quadratic response of the XC functional within the adiabatic approximation

$$\begin{aligned} v_{xc}^{xy}(\mathbf{r}, t) &= \int f_{xc}(\mathbf{r}, \mathbf{r}') \rho^{xy}(\mathbf{r}', t) d\mathbf{r}' \\ &+ \int \int g_{xc}(\mathbf{r}, \mathbf{r}', \mathbf{r}'') \rho^x(\mathbf{r}', t) d\rho^y(\mathbf{r}'', t) d\mathbf{r}' d\mathbf{r}'' \\ &= \int \frac{\partial^2 E_{xc}}{\partial \rho(\mathbf{r}) \partial \rho(\mathbf{r}')} \rho^{xy}(\mathbf{r}', t) d\mathbf{r}' \\ &+ \int \int \frac{\partial^3 E_{xc}}{\partial \rho(\mathbf{r}) \partial \rho(\mathbf{r}') \partial \rho(\mathbf{r}'')} \rho^x(\mathbf{r}', t) \rho^y(\mathbf{r}'', t) d\mathbf{r}' d\mathbf{r}'', \end{aligned} \quad (3.175)$$

and the resulting two-electron term in eq. 3.166 becomes

$$\mathbf{G}(\mathbf{P}^{xy}) \rightarrow \mathbf{G}_{DFT}(\mathbf{P}^{xy}, \mathbf{P}^x, \mathbf{P}^y) = \mathbf{J}(\mathbf{P}^{xy}) + \mathbf{V}_{xc}^{(2)}(\mathbf{P}^{xy}) + \mathbf{V}_{xc}^{(3)}(\mathbf{P}^x, \mathbf{P}^y), \quad (3.176)$$

where $\mathbf{V}_{xc}^{(2)}(\mathbf{P}^{xy})$ denotes the matrix representation of the $f_{xc}(\mathbf{r}, \mathbf{r}')$ part of $v_{xc}^{xy}(\mathbf{r}, t)$ and $\mathbf{V}_{xc}^{(3)}(\mathbf{P}^x, \mathbf{P}^y)$ the corresponding representation of the $g_{xc}(\mathbf{r}, \mathbf{r}', \mathbf{r}'')$. Note that the third order derivative of the XC energy is contracted with the first order transition densities only, so we can build the terms $\mathbf{V}_{xc}^{(3)}(\mathbf{P}^x, \mathbf{P}^y)$ before the iterative procedure starts.

In contrast to the second order derivatives of the XC energy functional, which has been implemented already in the context of TDDFT routines to determine excited states, the routines for the third order derivatives had to be implemented. Apart from the general routines, which are outlined in appendix A for LSDA and GGA functionals, the partial derivatives for some selected XC functionals have been implemented (see the table in appendix A).

3.7.2.3 Determination of Hyperpolarizabilities Exploiting Wigner's $(2n + 1)$ Rule

Following the $(2n + 1)$ rule of Wigner we can express the $(2n + 1)$ th order property using the n th order transition wavefunction or density, respectively. Corresponding formulas for MO-based schemes are available and are routinely used in most ab initio packages [86]. In this section we will derive analogous equations containing only matrix representations in the AO basis for the example of second harmonic generation $\beta_{xyz}(\mp 2\omega; \pm\omega, \pm\omega)$. Since we will start from the MO-based result, we index AO and MO quantities.

As can be seen from eq. 3.60, there are two types of terms containing either first order quantities represented in the MO basis. In order to derive an expression containing only matrices in the AO basis for the first group, the first term of eq. 3.60 is chosen as example

$$\begin{aligned} T_1 &= \text{Tr}_{\text{occ}} \left[\mathbf{U}^x(-2\omega) \mathbf{F}_{MO}^y(+\omega) \mathbf{U}^z(+\omega) \right] \\ &= \text{Tr} \left[\mathbf{P}^{MO} \mathbf{U}^x(-2\omega) \mathbf{C}^\dagger \mathbf{F}_{AO}^y(+\omega) \mathbf{C} \mathbf{U}^z(+\omega) \right], \end{aligned} \quad (3.177)$$

with \mathbf{P}^{MO} as diagonal matrix containing the occupation numbers of the single orbitals (see definition in eq. 3.61). Because of the idempotency of \mathbf{P}^{MO} ($= \mathbf{P}^{MO} \mathbf{P}^{MO}$) and the orthonormality of the MO coefficients we can insert $\mathbf{C}^\dagger \mathbf{S} \mathbf{C}$ and \mathbf{P}^{MO}

$$T_1 = \text{Tr} \left[\mathbf{C}^\dagger \mathbf{S} \mathbf{C} \mathbf{P}^{MO} \mathbf{U}^x(-2\omega) \mathbf{C}^\dagger \mathbf{F}_{AO}^y(+\omega) \mathbf{C} \mathbf{U}^z(+\omega) \mathbf{P}^{MO} \right]. \quad (3.178)$$

Since the trace is invariant to cyclic permutations we can partially reformulate this equation in terms of the density

$$T_1 = \text{Tr} \left[\mathbf{SCP}^{MO} \mathbf{U}^x(-2\omega) \mathbf{C}^\dagger \mathbf{F}_{AO}^y(+\omega) \mathbf{CU}^z(+\omega) \mathbf{P}^{MO} \mathbf{C}^\dagger \right], \quad (3.179)$$

where the diagonal matrix \mathbf{P}^{MO} is an intrinsic part of \mathbf{P} (eq. 2.36). Reconsidering the spectral definition of $\mathbf{P}^x(\pm\omega)$ in section 3.5, a multiplication of $\mathbf{P}^x(\pm\omega)$ with \mathbf{PS} from the left yields

$$\begin{aligned} \mathbf{PSP}^x(\pm\omega) &= \mathbf{CP}^{MO} \mathbf{C}^\dagger \mathbf{SC}^x(\pm\omega) \mathbf{P}^{MO} \mathbf{C}^\dagger + \mathbf{CP}^{MO} \mathbf{C}^\dagger \mathbf{SCP}^{MO} \mathbf{C}^{x\dagger}(\mp\omega) \\ &= \mathbf{CP}^{MO} \mathbf{C}^\dagger \mathbf{SCU}^x(\pm\omega) \mathbf{P}^{MO} \mathbf{C}^\dagger - \mathbf{CP}^{MO} \mathbf{C}^\dagger \mathbf{SCP}^{MO} \mathbf{U}^x(\pm\omega) \mathbf{C}^\dagger \\ &= \mathbf{CP}^{MO} \mathbf{U}^x(\pm\omega) \mathbf{P}^{MO} \mathbf{C}^\dagger - \mathbf{CP}^{MO} \mathbf{U}^x(\pm\omega) \mathbf{C}^\dagger \\ &= \mathbf{0} - \mathbf{CP}^{MO} \mathbf{U}^x(\pm\omega) \mathbf{C}^\dagger, \end{aligned} \quad (3.180)$$

where the first term vanishes because of the structure of \mathbf{U}^x (pattern on the right of Fig. 3.3). Since \mathbf{P}^{MO} has only non-zero elements in the occ/occ block, the symmetric projection of \mathbf{U}^x vanishes

$$\mathbf{P}^{MO} \mathbf{U}^x(\pm\omega) \mathbf{P}^{MO} = \mathbf{0}. \quad (3.181)$$

Analogously, we obtain by multiplying \mathbf{P}^x from the right with \mathbf{SP}

$$\begin{aligned} \mathbf{P}^x(\pm\omega) \mathbf{SP} &= \mathbf{C}^x(\pm\omega) \mathbf{P}^{MO} \mathbf{C}^\dagger \mathbf{SCP}^{MO} \mathbf{C}^\dagger + \mathbf{SCP}^{MO} \mathbf{C}^{x\dagger}(\mp\omega) \mathbf{SCP}^{MO} \mathbf{C}^\dagger \\ &= \mathbf{CU}^x(\pm\omega) \mathbf{P}^{MO} \mathbf{C}^\dagger \mathbf{SCP}^{MO} \mathbf{C}^\dagger - \mathbf{CP}^{MO} \mathbf{U}^x(\pm\omega) \mathbf{C}^\dagger \mathbf{SCP}^{MO} \mathbf{C}^\dagger \\ &= \mathbf{CU}^x(\pm\omega) \mathbf{P}^{MO} \mathbf{C}^\dagger - \mathbf{CP}^{MO} \mathbf{U}^x(\pm\omega) \mathbf{P}^{MO} \mathbf{C}^\dagger \\ &= \mathbf{CU}^x(\pm\omega) \mathbf{P}^{MO} \mathbf{C}^\dagger - \mathbf{0}. \end{aligned} \quad (3.182)$$

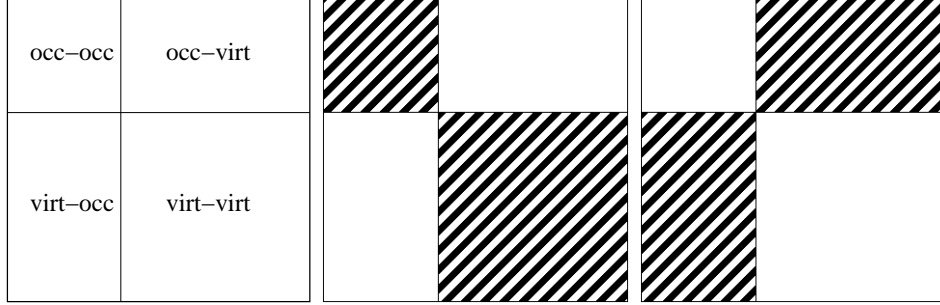
Inserting these expressions in eq. 3.179 one obtains a definition for T_1 exclusively using density matrices

$$\begin{aligned} T_1 &= -\text{Tr} \left[\mathbf{SPSP}^x(-2\omega) \mathbf{F}_{AO}^y(+\omega) \mathbf{P}^z(+\omega) \mathbf{SP} \right] \\ &= -\sum_{\mu\nu} S_{\mu\nu} [\mathbf{PSP}^x(-2\omega) \mathbf{F}_{AO}^y(+\omega) \mathbf{P}^z(+\omega) \mathbf{SP}]_{\mu\nu} \\ &= -\text{Dot} [\mathbf{S}, \mathbf{PSP}^x(-2\omega) \mathbf{F}_{AO}^y(+\omega) \mathbf{P}^z(+\omega) \mathbf{SP}]. \end{aligned} \quad (3.183)$$

For reformulating the second group of terms, we consider its first term as example:

$$T_2 = \text{Tr} \left[\mathbf{P}^{MO} \mathbf{U}^x(\mp 2\omega) \mathbf{U}^z(\pm\omega) \boldsymbol{\epsilon}^y(\pm\omega) \right]. \quad (3.184)$$

Figure 3.3: Matrix Patterns. Pattern of $\epsilon^x(\pm\omega)$ (middle) and $\mathbf{U}^x(\pm\omega)$ (right). White fields indicate zero entries.



First of all one has to analyze the expression for the first order energies (eq. 3.51)

$$\epsilon^y(\pm\omega) = \mathbf{F}_{\text{MO}}^y(\pm\omega) + \epsilon \mathbf{U}^y(\pm\omega) - \mathbf{U}^y(\pm\omega) \epsilon \pm \omega \mathbf{U}^y(\pm\omega). \quad (3.185)$$

By consideration of the patterns of the different matrices (Fig. 3.3) we will extract the significant part of $\epsilon^x(\pm\omega)$ in eq. 3.184. Since no non-Brillouin terms occur in a canonical solution, the orbital-energy matrix ϵ has a diagonal structure. A multiplication of an arbitrary matrix \mathbf{A} and a diagonal matrix yields a matrix that has the same structure as \mathbf{A} . Thus the last two terms of eq. 3.185 can be analyzed by the product of three transition coefficient matrices ($\mathbf{U}^i \mathbf{U}^j \mathbf{U}^k$) that all exhibit the pattern shown on the right of Fig. 3.3. In Fig. 3.4 the two subsequent multiplications are plotted schematically. While a first multiply yields an occ/occ-virt/virt pattern (upper scheme in Fig. 3.4), a further multiplication with a further \mathbf{U}^x matrix restores the original occ/virt-virt/occ structure.

Since all diagonal elements are zero the trace of such a matrix is also zero, thus we can replace

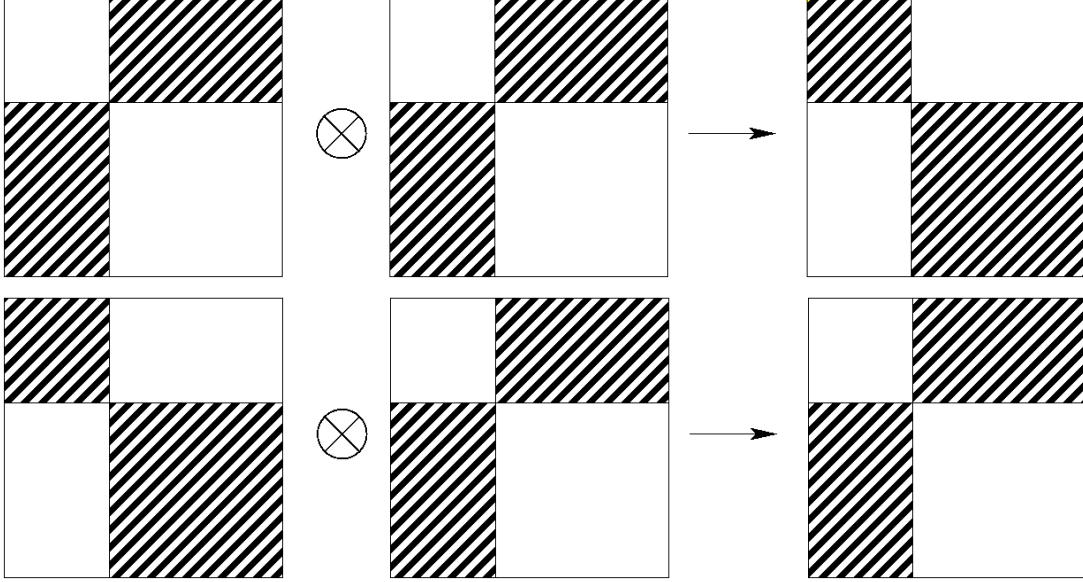
$$\epsilon^x(\pm\omega) \longrightarrow \mathbf{F}_{\text{MO}}^x(\pm\omega) \quad (3.186)$$

$$\implies T_2 = \text{Tr} \left[\mathbf{P}^{MO} \mathbf{U}^x(\mp 2\omega) \mathbf{U}^z(\pm\omega) \mathbf{F}_{MO}^y(+\omega) \right] \quad (3.187)$$

in eq. 3.184. Since the product of two transition coefficient matrices is block-diagonal, we can write

$$\mathbf{P}^{MO} \mathbf{U}^x(\mp 2\omega) \mathbf{U}^z(\pm\omega) = \mathbf{P}^{MO} \mathbf{U}^x(\mp 2\omega) \mathbf{U}^z(\pm\omega) \mathbf{P}^{MO}. \quad (3.188)$$

Figure 3.4: Multiplication of Matrix Patterns



Thus one can expand eq. 3.187 in a similar way as T_1

$$\begin{aligned}
 T_2 &= \text{Tr} \left[\mathbf{P}^{MO} \mathbf{U}^x (\mp 2\omega) \mathbf{U}^z (\pm \omega) \mathbf{P}^{MO} \mathbf{F}_{MO}^y (+\omega) \right] \\
 &= \text{Tr} \left[\mathbf{P}^{MO} \mathbf{U}^x (\mp 2\omega) \mathbf{C}^\dagger \mathbf{S} \mathbf{C} \mathbf{U}^z (\pm \omega) \mathbf{P}^{MO} \mathbf{C}^\dagger \mathbf{F}_{AO}^y (+\omega) \mathbf{C} \right] \\
 &= \text{Tr} \left[\mathbf{S} \mathbf{C} \mathbf{U}^z (\pm \omega) \mathbf{P}^{MO} \mathbf{C}^\dagger \mathbf{F}_{AO}^y (+\omega) \mathbf{C} \mathbf{P}^{MO} \mathbf{U}^x (\mp 2\omega) \mathbf{C}^\dagger \right]. \quad (3.189)
 \end{aligned}$$

Using eq. 3.180 and eq. 3.182 one obtains

$$\begin{aligned}
 T_2 &= -\text{Tr} \left[\mathbf{S} \mathbf{P}^x (-2\omega) \mathbf{S} \mathbf{P} \mathbf{F}_{AO}^z (+\omega) \mathbf{P} \mathbf{S} \mathbf{P}^y (+\omega) \right] \\
 &= -\text{Dot} \left[\mathbf{S}, \mathbf{P}^x (-2\omega) \mathbf{S} \mathbf{P} \mathbf{F}_{AO}^z (+\omega) \mathbf{P} \mathbf{S} \mathbf{P}^y (+\omega) \right]. \quad (3.190)
 \end{aligned}$$

So the final expression of the elements of the second harmonic generation tensor is

$$\begin{aligned}
 \beta_{xyz}(-2\omega; +\omega, +\omega) &= -\text{Dot}[\mathbf{S}, \mathbf{PSP}^x(-2\omega)\mathbf{F}_{AO}^y(+\omega)\mathbf{P}^z(+\omega)\mathbf{SP}] \\
 &- \text{Dot}[\mathbf{S}, \mathbf{PSP}^z(+\omega)\mathbf{F}_{AO}^y(+\omega)\mathbf{P}^x(-2\omega)\mathbf{SP}] \\
 &- \text{Dot}[\mathbf{S}, \mathbf{PSP}^x(-2\omega)\mathbf{F}_{AO}^z(+\omega)\mathbf{P}^y(+\omega)\mathbf{SP}] \\
 &- \text{Dot}[\mathbf{S}, \mathbf{PSP}^y(+\omega)\mathbf{F}_{AO}^z(+\omega)\mathbf{P}^x(-2\omega)\mathbf{SP}] \\
 &- \text{Dot}[\mathbf{S}, \mathbf{PSP}^z(+\omega)\mathbf{F}_{AO}^x(-2\omega)\mathbf{P}^y(+\omega)\mathbf{SP}] \\
 &- \text{Dot}[\mathbf{S}, \mathbf{PSP}^y(+\omega)\mathbf{F}_{AO}^x(-2\omega)\mathbf{P}^z(+\omega)\mathbf{SP}] \\
 &+ \text{Dot}[\mathbf{S}, \mathbf{P}^y(+\omega)\mathbf{SPF}_{AO}^x(-2\omega)\mathbf{PSP}^z(+\omega)] \\
 &+ \text{Dot}[\mathbf{S}, \mathbf{P}^z(+\omega)\mathbf{SPF}_{AO}^x(-2\omega)\mathbf{PSP}^y(+\omega)] \\
 &+ \text{Dot}[\mathbf{S}, \mathbf{P}^x(-2\omega)\mathbf{SPF}_{AO}^y(+\omega)\mathbf{PSP}^z(+\omega)] \\
 &+ \text{Dot}[\mathbf{S}, \mathbf{P}^z(+\omega)\mathbf{SPF}_{AO}^y(+\omega)\mathbf{PSP}^x(-2\omega)] \\
 &+ \text{Dot}[\mathbf{S}, \mathbf{P}^x(-2\omega)\mathbf{SPF}_{AO}^z(+\omega)\mathbf{PSP}^y(+\omega)] \\
 &+ \text{Dot}[\mathbf{S}, \mathbf{P}^y(+\omega)\mathbf{SPF}_{AO}^z(+\omega)\mathbf{PSP}^x(-2\omega)]. \tag{3.191}
 \end{aligned}$$

For TDDFT calculations the XC kernel g_{xc} , that contains the third order derivatives of the XC energy functional, has to be considered [118] by adding the term

$$\text{Dot}[\mathbf{P}^x(-2\omega), \mathbf{V}_{xc}^{(3)}(\mathbf{P}^y(+\omega), \mathbf{P}^z(+\omega))]. \tag{3.192}$$

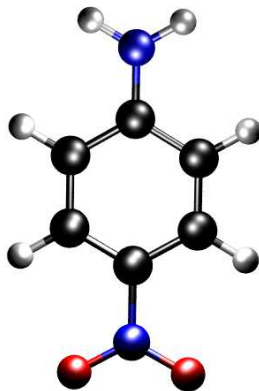
Since the expression in eq. 3.191 and eq. 3.192 only contains linear scaling matrices (i.e. the number of significant elements grows proportional with the system size), we can apply sparse algebra routines to obtain the result in $\mathcal{O}(M)$ fashion.

Note that this approach has several advantages compared to a quadratic response calculation. While the latter needs nine TDSCF calculations to obtain the quadratic response – where the convergence of the second order TDSCF also clearly depends on the quality of the first order result $\mathbf{P}^x(\pm\omega)$ – we only have to perform six first order TDSCF calculations (for $\pm\omega$ and $\pm 2\omega$) within the Wigner approach.

3.7.2.4 First Molecular Hyperpolarizabilities – Illustrative Examples

The non-linear optical properties of a crystal depend on the different components (ions, molecules) and the interaction between them. In contrast to inorganic materials, whose

Figure 3.5: p-Nitroaniline



non-linear properties mainly originate from lattice vibrations because of the strong interactions between the different components, the properties of weak-interacting organic crystals mainly depend on the high mobility of the delocalized π -electrons. Since the effects in organic materials are evidently faster than that of inorganic compounds, the interest in organic materials for optical devices has increased in the last years. A first glance at the efficiency of such devices like optical fibres, optical frequency converters, electro-optic modulators, thermo-optic switches etc. but also liquid crystals for TFT displays is given by the polarizabilities and hyperpolarizabilities [26].

The scaling behavior of the calculation of the third order properties using Wigner's $(2n+1)$ rule is shown for the example of a series of linear alkanes at the TDDFT-BP86(VWN)/6-31G* [109–111] level of theory in Fig. B.8. Since the calculations are based on the first order results given in section 3.7.1.2, only the timings for the formation of $\mathbf{V}_{xc}^{(3)}(\mathbf{P}^x, \mathbf{P}^y)$ and the contractions in eq. 3.191 are given in Fig. B.8. In order to be consistent with the first order calculations, a sparse algebra threshold of $\text{thr}_{\text{SA}} = 10^{-7}$ is used. Since the formation of the third order derivatives of the XC energy functional scales perfectly linear with system size, the $\mathcal{O}(M^3)$ multiplications (cubes) to form eq. 3.191 clearly dominate the overall computational time. The application of sparse multiplications (circles) shows an $\mathcal{O}(M)$ scaling behavior that becomes superior to the standard multiplications at approx. 2400 basis functions.

As an example for the accuracy of the presented method the calculation of para-nitroaniline (PNA, Fig. 3.5) at HF and KS-DFT level is presented. The calculation is

Table 3.6: Second harmonic generation for PNA with B3LYP/6-31G* with different frequencies ω . Average values $\bar{\beta}_z$ using the B-convention [119] are given in 10^3 esu.

Freq. [eV]	This Work (Wigner)	Ref. [120]	Ref [121] (Exp.)
0.650	6.78	6.50	9.60±0.5
1.170	10.11	9.50	16.9 ±0.4
1.361	12.96	12.00	25±1
1.494	16.25	14.84	40±3

done by exploiting Wigner’s $(2n + 1)$ rule based on the dynamic polarizabilities presented in section 3.7.1.2. Because of the amino and nitro groups at the phenyl ring this system exhibits a strong charge-transfer character and has become a standard test system for theoretical and experimental methods [118–120]. In all calculations the structure with C_{2v} symmetry is used (data is given in Tab. C.1). It has been obtained from a geometry optimization with B3LYP/6-31G* [115, 116]. Note that the B-convention [119] is used

$$\bar{\beta}_i = \frac{1}{3} \sum_j (\beta_{ijj} + \beta_{iji} + \beta_{jji}). \quad (3.193)$$

The different first hyperpolarizabilities are calculated at the TDHF and TDDFT/B3LYP level with the 6-31G* basis set. The calculations were performed with a tight convergence criterion and an integral threshold of 10^{-10} . As numerical grid for the DFT quadrature a Euler-McLaurin/Lebedev-Laikov (75,302) grid is chosen [122–124]. The data for different frequencies ω in tables C.4 and C.5 show that the new density matrix-based scheme yields the same results as the traditional MO-based method. Furthermore, it has to be mentioned that the results obtained by a second order D-TDDFT calculation, i.e. without using Wigner’s rule, are of course the same. The following comparison of the SHG value $\bar{\beta}_z$ (in 10^3 esu) at different frequencies ω calculated with B3LYP/6-31G* shows that the results are close to those given in Ref. [120]. The remaining deviations might be due to differences in the structural parameters. In order to get a first estimate for those influences, we stretched just one bond (between the amino group and the phenyl ring) by 0.01 Å. The modified structure shows a change in $\bar{\beta}_z(-2\omega; +\omega, +\omega)$ by $+0.08 \cdot 10^3$ esu (for $\omega = 0.650$ eV). Thus one can state that there is a strong influence on the elements of the hyperpolarizability tensor and so the difference observed is expected to be due to the different structure

compared to the work of Salek et al. [120]. In agreement with the work of van Gisbergen et al. [118], we also found only a small impact of the term in eq. 3.192, which contains the third order derivative of the XC energy functional, on β (max. 1.9%).

Additionally to the SHG the static hyperpolarizability and the electro-optical Pockel's effect with $\omega = 0.650$ eV has been calculated. We found the same order in magnitudes on HF and B3LYP level as in Ref. [119]

$$\bar{\beta}_z(0; 0, 0) > \bar{\beta}_z(0; +\omega, -\omega) > \bar{\beta}_z(-2\omega; +\omega, +\omega),$$

HF:	3.80	>	3.91	>	4.15,
B3LYP:	5.81	>	6.11	>	6.78.

Note that the B-convention gives the same results for the optical rectification and the electro-optical Pockel's effect, i.e. $\bar{\beta}_z(-\omega; 0, +\omega) = \bar{\beta}_z(0; +\omega, -\omega)$, where the two quantities are related as follows [119]:

$$\begin{aligned}\beta_{xxz}(0; +\omega, -\omega) &= \beta_{zxx}(-\omega; 0, +\omega), \\ \beta_{yyz}(0; +\omega, -\omega) &= \beta_{zyy}(-\omega; 0, +\omega), \\ \beta_{zzz}(0; +\omega, -\omega) &= \beta_{zzz}(-\omega; 0, +\omega).\end{aligned}$$

Chapter 4

Analysis of Sparse Algebra Routines within the D-GIAO-HF Algorithm

We implemented a C++ class to handle $N \times N$ matrices in sparse format via the row indexed sparse storage mode (RISSM) as described by Press et al. [105], which is a slight modification of the compressed sparse row (CSR) scheme [125]. Here only the significant values of the matrix relative to a given sparse algebra (SA) threshold thr_{SA} are stored.

The crucial part of the presented density matrix-based methods is the matrix multiplication ($\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$) that clearly dominates the algebraic routines because of the $\mathcal{O}(M^3)$ scaling behavior. In this work several routines to perform these multiplications have been implemented that are similar to the algorithm presented by Gustavson [125]. The basic structure of these routines (scheme in Fig. B.14) as well as the different screening schemes are described in section 4.1. Note that the presented exemplary applications of the previous chapter are calculated with the unscreened multiplication algorithm (SMT-x ($x=\infty$)).

In section 4.4 the performance of the sparse multiplication routines is compared with an $\mathcal{O}(M^3)$ routine of a standard linear algebra library. For small systems the $\mathcal{O}(M^3)$ multiplications of the Intel Math Kernel library (MKL) [126] are clearly faster than our $\mathcal{O}(M)$ sparse routines because of the overhead (factor approx. 16-18 for the multiplications in sec. 4.4) caused by the index controlled access of matrix elements, but with increasing system size our sparse routines become superior to the $\mathcal{O}(M^3)$ routine. The homologous series of linear alkanes with (GIAO-HF/6-31G*) shows a crossover at approx. 1700 basis functions^{a)}.

^{a)}See section 3.6.4

As mentioned in section 3.6.1.3, the linear scaling behavior is hampered by the decrease of sparsity after the transformation of the residual matrix from the co- into the contravariant basis. Other multiplications like $\mathbf{P} \times \mathbf{S}$ change the number of significant elements compared to the use of $\mathcal{O}(M^3)$ matrix multiplications only slightly. The reason can be found in the truncation of matrices and maximal absolute values which are larger for the inverse metric \mathbf{S}^{-1} by at least an order of magnitude. These effects are investigated in section 4.3. In order to minimize the number of multiplications with \mathbf{S}^{-1} we use the contravariant LEQS in eq. 3.105, where — in combination with the pre-built matrices \mathbf{FS}^{-1} and \mathbf{PS} — we only need four multiplications to build the LHS on Level-1.

4.1 Screening in Sparse Matrix Multiplications

As it has been mentioned, the sparse multiplication routines are similar to the algorithm proposed by Gustavson [125]. Therefore the outer loop is over the rows of matrix \mathbf{A} as it is depicted in Fig. B.14. For the given row i of \mathbf{A} the corresponding row i of the resulting matrix \mathbf{C} is built incrementally by multiplying each element A_{ij} with all elements of row j of matrix \mathbf{B} . The routines provide basically the same structure and only differ in the definition of the effective threshold thr_{eff} used for the internal screening. The screening procedure is similar to the pre-ordering in the LinK method [8], i.e. the rows of the matrix on the right (matrix \mathbf{B}) are sorted in decreasing order with a `quicksort` algorithm, so that the innermost loop in Fig. B.14 is stopped, once the absolute value of the actual element in the row of matrix \mathbf{B} is smaller than the threshold thr_{eff} . It has to be mentioned that the `quicksort` routine scales with $\mathcal{O}(M^2)$ in the worst case for a vector of length M [105], but as long as just the rows of a linear scaling matrix are ordered – since it has an average, constant number of significant elements per row – the sorting of a single row scales with $\mathcal{O}(\text{const}^2)$. In the following the routines are abbreviated with the prefix "SMT-" (sparse modified thresholding) and the endings A, B, C, D and x.

Prior to the discussion of the validity and performance of the presented SMT modifications, the rigorous screening within a typical $\mathcal{O}(M^3)$ multiplication procedure is discussed, i.e. row times column of $N \times N$ matrices

$$C_{ij} = \sum_k^{n_{\text{prod}} \leq N} A_{ik} B_{kj}, \quad (4.1)$$

where n_{prod} is the number of products of significant elements $|A_{ik}|, |B_{kj}| \geq \text{thr}_{SA}$. Splitting the sum into products with absolute values larger than or equal to an effective threshold $\text{thr}_{eff} = \text{thr}_{SA}/n_{prod}$ (sig) and products smaller than thr_{eff} (unsig) yields

$$C_{ij} = \sum_k^{n_{sig}} A_{ik} B_{kj} + \sum_k^{n_{unsig}} A_{ik} B_{kj}. \quad (4.2)$$

By discarding the second sum

$$C_{ij} \approx \sum_k^{n_{sig}} A_{ik} B_{kj}, \quad (4.3)$$

the error in C_{ij} remains smaller than thr_{SA} since

$$\begin{aligned} n_{unsig} &\leq n_{prod}, \\ \sum_k^{n_{unsig}} A_{ik} B_{kj} &< n_{prod} \cdot \text{thr}_{eff} = \text{thr}_{SA}. \end{aligned} \quad (4.4)$$

Reconsidering the algorithm in Fig. B.14 one can see that this rigorous screening is not feasible^{b)} since we would have to order the rows of matrix \mathbf{B} with respect to two criteria. First, the absolute value of the different products $A_{ik} B_{kj}$, and second, the effective threshold for the given element C_{ij} . Intuitively, one may replace n_{prod} with N to form the effective threshold $\text{thr}_{eff} = \text{thr}_{SA}/N$ still providing a valid lower bound. It is easily seen that this approach is not size-consistent and also becomes ineffective with increasing N ($\lim_{N \rightarrow \infty} \text{thr}_{eff} = 0 \rightarrow \text{SMT-x}$ ($x = \infty$)). The schemes denoted as SMT-A, SMT-B and SMT-C all provide valid lower bounds to thr_{eff} by using different approximated values \tilde{n}_{prod} for n_{prod} with respect to the i th row A_i of \mathbf{A} .

$$\begin{aligned} \text{SMT-A:} \quad \tilde{n}_{prod} = \tilde{n}_{AB1} &= \min(\text{length}(A_i), \max(\text{lengths of columns of significant} \\ &\quad \text{elements in actual row of } \mathbf{B})) \end{aligned} \quad (4.5)$$

$$\text{SMT-B:} \quad \tilde{n}_{prod} = \tilde{n}_{AB2} = \min(\text{length}(A_i), \max(\text{lengths of all columns of } \mathbf{B})) \quad (4.6)$$

$$\text{SMT-C:} \quad \tilde{n}_{prod} = \tilde{n}_{AB3} = \min(\text{length}(A_i), \text{average length of columns of } \mathbf{B}) \quad (4.7)$$

$$\text{SMT-D:} \quad \tilde{n}_{prod} = \tilde{n}_A = \text{length}(A_i) \quad (4.8)$$

The overhead of the determination of \tilde{n}_{prod} but also the efficiency of the screening process decreases from SMT-A to SMT-D as can be seen from the example of an amylose

^{b)}Except for the case that we already know which values of the resulting matrix \mathbf{C} are significant, but this would still exhibit a sensible overhead.

Table 4.1: CPU timings for 38 Level-1 iterations of an amylose chain containing 8 α -D-glucose units with different sparse multiplication screening schemes. The sparsity refers to a converged perturbed density matrix \mathbf{P}^x .

Screening	CPU [min]	MegaFlops (1st Level-1)	Sparsity [%]
SMT-x ($x=\infty$)	66.85	18191	14.5
SMT-A	70.65	15191*	14.5
SMT-B	70.96	15169	14.5
SMT-C	78.24	18075	14.4
SMT-D	77.11	18213	14.5

*15166 MegaFlops if same overhead as SMT-B.

chain containing 8 α -D-glucose units depicted in Fig. B.12, the CPU timings for the sum of all 38 Level-1 iterations can be seen from Tab. 4.1. We see that the SMT-A screening is the most effective, but the crossover with the unscreened routine has not taken place for this size of system. The contradiction between the times and flop-counts of SMT-A and SMT-B in Tab. 4.1 and Fig. B.12 results from the larger overhead of SMT-A. While less $A_{ik}B_{kj}$ multiplies (smaller contraction length in innermost loop) are needed, the SMT-A screening requires a further multiplication (see remark in Tab. 4.1).

Since the crossover depends on the system as well as the chosen basis set we used unscreened multiplications in the calculations of the illustrative examples in section 3.6.4. However, the intrinsic compression of the resulting matrix could also be understood as a succeeding screening procedure.

Note that the presented SMT-x modification is not meant as real alternative to the other algorithms, it should only show the huge deteriorating impact on the results if invalid screening schemes are used (see section 4.2).

4.2 Behavior and Stability of Sparse Algebra Routines within an Iterative Process

In section 3.6.1.3 it has been mentioned that the occ/occ part from \mathbf{P}_{ov}^x differs from the expression one obtains from the derivative of the idempotency condition ($-\mathbf{P}\mathbf{S}^x\mathbf{P}$) if sparse

algebra routines are applied (see eq. 3.100)

$$\Delta \mathbf{P}_{\text{oo}}^x = \mathbf{P}_{\text{oo}}^x - \mathbf{P}_{\text{oo}}^{x'} \neq \mathbf{0}. \quad (4.9)$$

The effect of approximate, i.e. truncated, matrices is shown for the example of this difference $\Delta \mathbf{P}_{\text{oo}}^x$ which vanishes if $\mathcal{O}(M^3)$ multiplications are used. In Tab. 4.2 the number of significant elements N_{sigel} is shown as a soft criterion to compare the matrices. In the column containing the value of the largest element (max. el.) the value is replaced by the norm if formally a zero matrix is expected. The difference $\Delta \mathbf{P}_{\text{oo}}^x$ is shown in the fifth and sixth row, respectively. Note that $N_{\text{sigel}}(\Delta \mathbf{P}_{\text{oo}}^x) = 1428$ is evidently larger than the difference ($N_{\text{sigel}}(\mathbf{P}_{\text{oo}}^{x'}) - N_{\text{sigel}}(\mathbf{P}^x \mathbf{S} \mathbf{P}) = 384$) since the truncation of matrices also falsifies larger values around the given threshold that cancel each other if the matrices are treated exactly (see sec. 4.3). The last entry in Tab. 4.2 shows the cancellation of the deviating $\mathbf{P}_{\text{oo}}^{x'}$. Note that this cancellation is intrinsic to our D-CPSCF equations in eq. 3.98 and eq. 3.105 in contrast to the alternative approaches described in section 3.6.1.4.

Additionally, the deficiencies of an improper screening scheme are presented. The results of the SMT-0 ($x=0$) multiplications in Tab. 4.2 show a large amount of numerical noise. At the example of $\mathbf{P} \mathbf{S} \mathbf{P}^x \mathbf{S} \mathbf{P} - \mathbf{P} \mathbf{S} (\mathbf{P} \mathbf{S} \mathbf{P}^x)^\dagger$ we can also see that the order of matrices in the multiplication affects the result (compared to $((\mathbf{P} \mathbf{S} \mathbf{P}^x)_{\text{oo}} - (\mathbf{P}^x \mathbf{S} \mathbf{P})_{\text{oo}})$) clearly showing the insufficiency of the screening scheme.

The stability of sparse multiplications within an iterative scheme has been tested by the multiple application of the purification transformation by McWeeny [47] to the unperturbed density $\mathbf{P} = 3\mathbf{P} \mathbf{S} \mathbf{P} - 2\mathbf{P} \mathbf{S} \mathbf{P} \mathbf{S} \mathbf{P}$ as is shown in Figs. B.9-B.11. In Fig. B.9 the largest absolute value of \mathbf{P} is plotted for each purification step. While the SMT-0 screening even strongly affects the largest elements, the unscreened multiplications show a stable behavior. It has to be mentioned that the sparsity is closely connected to the accuracy of the multiplication which can be seen in Fig. B.11. The sparsity is not or – at least – only slightly affected when using the unscreened or the balanced SMT-A/B/C algorithms, otherwise it decreases constantly resulting in a dense matrix which hampers efficient sparse routines.

For the model system HCF_3 (HF/6-31G*) we have tested the modified routines using effective thresholds thr_{eff} outlined in Fig. B.14. Except for SMT-x with $x=0$ the modified routines obtain stable results with respect to the change in the maximal element (Fig. B.9) as well as the sparsity of \mathbf{P} (Fig. B.11).

Table 4.2: HCF₃, HF/6-31G* (N_{BF}=62, N_{BF}²=3844). Analysis of occ/occ part of \mathbf{P}_{ov}^x and \mathbf{P}_{vo}^x and multiplication order of screened and unscreened sparse multiplications. All values (N_{sigel}: number of significant elements; max. el.: value of elements with largest absolute value) with respect to the threshold 10^{-8} (as soft criterion for comparison). \mathbf{P}^x is taken from the last CPSCF iteration after DIIS extrapolation.

Matrix	not screened		screened/SMT0		screened/SMT-A		$\mathcal{O}(M^3)$
	N _{sigel}	max. el.	N _{sigel}	max. el.	N _{sigel}	max. el.	N _{sigel}
$(\mathbf{PSP}^x)_{oo}$	3409	9.32e-02	3637	9.32e-02	3458	9.32e-02	3026
$(\mathbf{P}^x\mathbf{SP})_{oo}$	3409	9.32e-02	3643	9.32e-02	3458	9.32e-02	3026
$(\mathbf{PSP}^x)_{oo} - (\mathbf{P}^x\mathbf{SP})_{oo}$	23	8.97e-10*	1238	1.11e-08*	24	9.60e-10*	0
$\mathbf{PS}^x\mathbf{P}$	3025	9.32e-02	3046	9.32e-02	3025	9.32e-02	3026
$(\mathbf{PSP}^x)_{oo} + \mathbf{PS}^x\mathbf{P}$	1428	1.20e-07*	2958	2.94e-06*	1505	2.31e-07*	0
$(\mathbf{P}^x\mathbf{SP})_{oo} + \mathbf{PS}^x\mathbf{P}$	1431	1.20e-07*	2956	2.94e-06*	1503	2.31e-07*	0
$\mathbf{PSP}^x\mathbf{SP} - \mathbf{PS}(\mathbf{PSP}^x)^\dagger$	23	8.97e-10*	1275	1.17e-08*	24	9.60e-10*	0
$(\mathbf{PSP}^x - \mathbf{P}^x\mathbf{SP})_{oo}$	5	5.02e-10*	845	9.16e-09*	4	4.46e-10*	0

*Norm of (difference) vector ($\|\mathbf{A}\| = [1/N_{BF}^2 \sum_{ij} |A_{ij}|^2]^{1/2}$).

4.3 The Influence of Truncating Matrices

Even if no screening in matrix multiplications is applied (SMT-x (x=∞) in Fig. B.14), there is still an intrinsic error emerging from the compression of the matrices with respect to the sparse algebra threshold thr_{SA} . The resulting deviation from the exact result using $\mathcal{O}(M^3)$ routines is traced by comparison of the formation of \mathbf{PS} and $\mathbf{S}^{-1}\mathbf{F}$ (Tab. 4.3 and Tab. 4.4), both appearing in our LEQS eq. 3.98 and eq. 3.105, respectively.

We suggest a simple guess to the order of the deviation $\mathcal{O}(\text{dev})$

$$\mathcal{O}(\text{dev}) = (\text{MAV}[\mathbf{A}] + \text{MAV}[\mathbf{B}]) \text{thr}_{SA}, \quad (4.10)$$

where $\text{MAV}[\mathbf{A}]$ is the maximal absolute value of matrix \mathbf{A} . Here the sum and not the average is our quantity of choice since the truncation of both matrices affect the result. The error results from the missing contributions of elements larger than 1 with elements smaller

Table 4.3: Effect of truncating matrices for several systems (HF/6-31G*), $\text{thr}_{\text{SA}} = 10^{-8}$ in the example of calculating $\mathbf{S}^{-1}\mathbf{F}$. The largest absolute matrix element is given while the index "diff" is the difference between the dense and the sparse result (SMT-x(x= ∞)). For a description of the guess to the order of the truncation error see the text.

System	\mathbf{S}^{-1}	\mathbf{F}	$\mathbf{S}^{-1}\mathbf{F}_{\text{diff}}$	$\mathcal{O}(\text{dev})$	$\mathbf{S}^{-1}\mathbf{F}$	\mathbf{L}^{-1}
C ₅ H ₁₂	1.199e+02	1.121e+01	1.115e-06	1.311e-06	1.117e+01	4.204e+00
C ₁₀ H ₂₂	1.520e+02	1.121e+01	3.512e-06	1.632e-06	1.117e+01	4.289e+00
C ₂₀ H ₄₂	1.535e+02	1.121e+01	3.674e-06	1.648e-06	1.117e+01	4.212e+00
C ₄₀ H ₈₂	1.536e+02	1.121e+01	3.452e-06	1.648e-06	1.117e+01	4.198e+00
Toluene	3.854e+02	1.123e+01	3.141e-07	3.966e-06	1.118e+01	1.026e+01
TMS	1.086e+02	6.874e+01	2.175e-07	1.774e-06	6.853e+01	7.762e+00

than thr_{SA} . Considering the MaxEl of \mathbf{S}^{-1} of toluene ($3.85377 \cdot 10^2$) and $\text{thr}_{\text{SA}} = 10^{-8}$ for example, the product with an insignificant element of \mathbf{F} with value $9 \cdot 10^{-9}$ would contribute $3.5 \cdot 10^{-6}$ to an element of $\mathbf{S}^{-1}\mathbf{F}$.

As mentioned before, the matrix products with the inverse metric generate relatively large errors due to large MaxEls compared to the other matrices (Tab. 4.3). In a D-

Table 4.4: Effect of truncating matrices for several systems (HF/6-31G*), $\text{thr}_{\text{SA}} = 10^{-8}$ at the example of calculating \mathbf{PS} . The largest absolute matrix element is given while the index "diff" is the difference between the dense and the sparse result (SMT-x(x= ∞)). For a description of the guess to the order of the truncation error see the text.

System	\mathbf{P}	\mathbf{S}	$\mathbf{PS}_{\text{diff}}$	$\mathcal{O}(\text{dev})$	\mathbf{PS}
C ₅ H ₁₂	1.030e+00	1.000e+00	2.263e-08	2.030e-08	9.984e-01
C ₁₀ H ₂₂	1.030e+00	1.000e+00	2.553e-08	2.030e-08	9.985e-01
C ₂₀ H ₄₂	1.030e+00	1.000e+00	2.564e-08	2.030e-08	9.985e-01
C ₄₀ H ₈₂	1.030e+00	1.000e+00	2.808e-08	2.030e-08	9.985e-01
Toluene	1.030e+00	1.000e+00	5.199e-09	2.030e-08	9.983e-01
TMS	1.127e+00	1.000e+00	3.675e-09	2.127e-08	9.995e-01

CPSCF calculation, where the transformation in eq. 3.102 or eq. 3.103 is multiply used, these errors also strongly decrease the sparsity hampering the efficiency of the sparse algebra routines. The reduction of the number of \mathbf{S}^{-1} multiplications and its implicit use via $\mathbf{S}^{-1}\mathbf{F}^c$) by solving the contravariant LEQS in eq. 3.105 reduce this negative effect. Alternatively the explicit use of \mathbf{S}^{-1} can be circumvented by the inverse Cholesky-factor \mathbf{L}^{-1} ($\mathbf{S}^{-1} = \mathbf{L}^{-1}\mathbf{L}^{-1\dagger}$) that not only has a smaller MaxEl (Tab. 4.3) but also provides a route to its linear scaling formation [127]. However, the use of the Cholesky factor requires a further multiplication compared to the presented approach using the pre-built matrix $\tilde{\mathbf{F}} = \mathbf{S}^{-1}\mathbf{F}$.

4.4 Comparison with Standard Library Routines

In this section the performance of the sparse multiplication routine $\mathbf{AB} = \mathbf{C}$ without screening is compared to the $\mathcal{O}(M^3)$ routine `dgemm` from the Intel Math Kernel Library [126], which is in our opinion the fastest available one. All calculations were done on an Intel Xeon EM64T architecture (64bit) using one 3.0GHz processor and the Intel C compiler [128] on a Linux system.

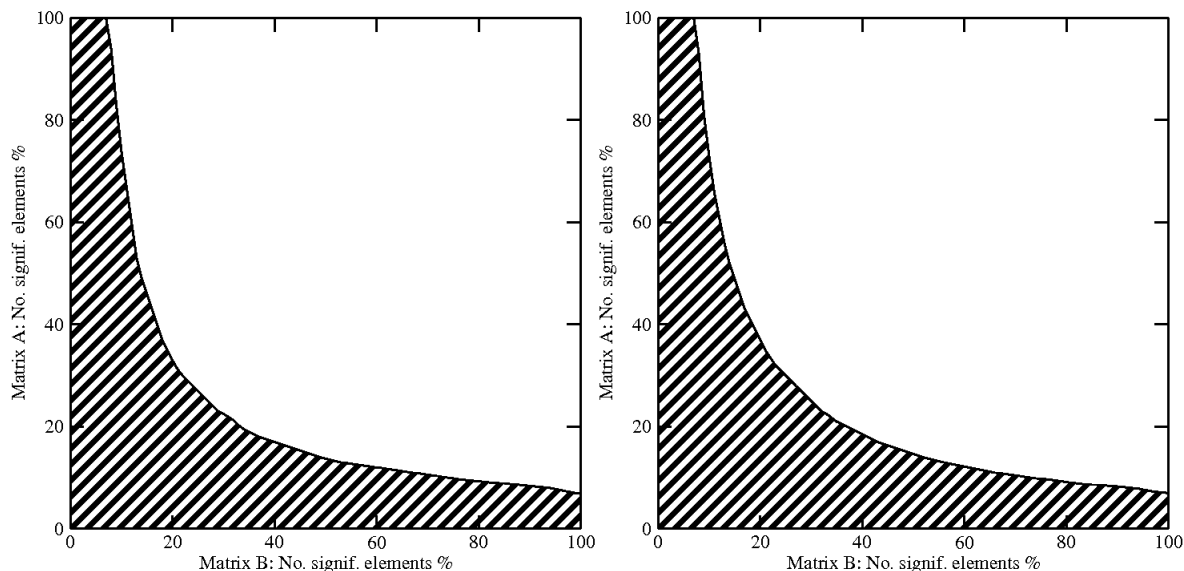
Table 4.5: Timings (CPU [s]) of sparse multiplications for a fixed sparsity of matrix \mathbf{A} (20%). $\mathcal{O}(M^3)$ times[126]: 0.42s for 1000x1000, 10.88s for 3000x3000 and 50.30s for 5000x5000.

Dimension	Sparsity in percent of matrix \mathbf{B}									
	1	5	10	15	20	25	30	34	36	38
1000	0.03	0.07	0.13	0.19	0.25	0.30	0.37	$> \mathcal{O}(M^3)$		
3000	0.40	1.69	3.24	4.74	6.28	7.71	9.22	10.37	$> \mathcal{O}(M^3)$	
5000	1.65	7.31	14.13	20.92	27.53	34.17	40.77	45.95	48.53	$> \mathcal{O}(M^3)$

In order to analyze the performance of the sparse multiplication routine, different combinations of matrices, each with a sparsity between 1% and 100%, are multiplied. Since the general performance of the routine should be tested, i.e. independent on the structure

^{c)}Note the smaller MaxEl of $\mathbf{S}^{-1}\mathbf{F}$ compared to \mathbf{S}^{-1} in Tab. 4.3.

Figure 4.1: Performance of sparse matrix multiplications $\mathbf{AB} = \mathbf{C}$ ($\text{thr}_{\text{SA}} = 10^{-7}$): Dimension 1000x1000 (left, $\mathcal{O}(M^3)$: 0.42s) and 5000x5000 (right, $\mathcal{O}(M^3)$: 50.3s), hatched areas show regions where sparse multiplications are faster than the $\mathcal{O}(M^3)$ routine [126]. The axes show the percentage of significant elements of matrices \mathbf{A} and \mathbf{B} .



of the matrices^{d)} (band diagonal, block tridiagonal etc.), the position of the significant elements as well as their values (abs. values between $10^1 - 10^{-7}$) are created by a pseudo random number generator [105]. In Fig. 4.1 the y-axis shows the number of significant elements of matrix \mathbf{A} in percent and the x-axis the one of matrix \mathbf{B} . The combinations of matrices in the hatched area are faster than the $\mathcal{O}(M^3)$ routine.

In Tab. 4.5 we extracted one dimension of the plots with a fixed sparsity of \mathbf{A} at 20%. As can be seen, the winning area grows with an increasing leading dimension N of the matrices. Furthermore, the sparsity of a linear scaling matrix grows with increasing molecular size M ($\lim_{N \rightarrow \infty} N_{\text{sigel}} = 0\%$). The two effects ensure that the efficiency of the sparse multiplication routine grows with increasing system size. Nevertheless, the onset of the linear scaling behavior clearly depends on the localization of the electrons in the chosen system (see examples). There is no doubt that with increasing computer power and simultaneously increasing system sizes the use of sparse algebra routines becomes more and more important.

^{d)}The sparse multiplication routines are also used in the QMC algorithms presented in chapter 5, where the occurring matrices can exhibit arbitrary structures.

Chapter 5

Linear Scaling Quantum Monte Carlo Algorithms for the Local Energy with Rigorously Controllable Error Bounds

The treatment of large molecular systems at the level of Hartree-Fock (HF) or Kohn-Sham density functional theory (KS-DFT) is routine nowadays [38]. However, the application of high-accurate post-Hartree-Fock correlation methods to obtain more reliable results are hampered by their unfavorable scaling behavior, e.g. $\mathcal{O}(N^7)$ for CCSD(T). It is obvious that for larger systems suitable methods need to scale as $\mathcal{O}(N-N^3)$, depending on the local or non-local nature of the electronic structure.

While quantum Monte Carlo methods are well established in the solid-state physics community, a growing interest of theoretical chemists in quantum Monte Carlo methods was noticeable in recent years. Using trial wave functions of Slater-Jastrow type Ψ_{SJ}^T the most commonly used variational (VQMC) [129] and fixed-node diffusion (FN-DQMC) [130, 131] quantum Monte Carlo methods scale cubically with the number of electrons for the local energy E_{loc} while e.g. FN-DQMC provides typically CCSD(T) accuracy [132]. Recent advances [133] based on the local approximation proposed by Pulay [134] induced several researchers to develop $\mathcal{O}(N)$ FN-DQMC methods [135–139] for a fixed sample size using localized molecular orbitals (LMO). Furthermore the introduction of local correlation factors of Boys-Handy type [140, 141] by Manten and Lüchow [135] provides a possible

early onset of an $\mathcal{O}(N)$ behavior. It has to be noted that the LMO approach introduces an empirical, distance-based cut-off parameter that results in an additional source of errors in MO based correlation methods. This parameter clearly depends on the system under investigation and always has to be carefully converged with the resulting energy.

The method presented in this section avoids the explicit use of MOs by reformulating the VQMC scheme in the basis of the N -particle density matrix emerging from a single Slater determinant. Transferring the spirit of e.g. density matrix-based SCF schemes to VQMC enables a linear scaling behavior by exploitation of the sparsity of the occurring quantities while preserving accurate results within the given error bounds.

After a brief review of the stochastic foundations of Monte Carlo methods and the popular VQMC and DQMC algorithms, the basic equations of the new N -PDM VQMC will be derived. Finally, a new approach to the DQMC method in terms of the N -particle density matrix is described. Since this new N -PDM DQMC algorithm allows to exploit the permutational symmetry information of the fermionic system, its application within the fixed-node approximation becomes possible. The scaling behavior of these new algorithms as implemented in our QUANTUMMC [142] package, which has been developed in this work, is shown for the example of a series of linear alkanes.

5.1 Stochastic Foundations

Within the Monte Carlo schemes one solves the n dimensional integral

$$I = \int f(\mathbf{r})d\mathbf{r}, \quad (5.1)$$

with stochastical methods. Standard numerical integration algorithms like the Newton-Cotes, the Simpson formula, or the Gauss quadratures [105] are well proven to obtain accurate results for low-dimensional functions. Since their error bounds are proportional to $L^{-c/n}$ with L as the number of integration points and c as constant depending on the quadrature scheme, higher dimensional functions ($n \geq 8$) would require an enormous number of sampling points. As will be shown later, the error bound of the Monte Carlo integration is proportional to $L^{-1/2}$, i.e. it is independent of dimensionality n . Therefore MC methods have become the standard methods for the integration of high-dimensional functions.

Assuming a set of L configurations $\{\mathbf{R}_i\}$, also called random walkers, distributed according to the normalized probability distribution $\mathcal{P}(\mathbf{R})$, the Monte Carlo estimate of the integral in eq. 5.1 is the average

$$\begin{aligned}\langle I \rangle &= \int_{-\infty}^{\infty} f(\mathbf{R})\mathcal{P}(\mathbf{R})d\mathbf{R} \\ &= L^{-1} \sum_{i=1}^L f(\mathbf{R}_i),\end{aligned}\tag{5.2}$$

converging to the exact result in the limit $L \rightarrow \infty$ according to the law of large numbers [143]. In order to obtain a reasonable estimate for the accuracy of the average $\langle I \rangle$, we have to impose some constraints on the set of configurations and the probability $\mathcal{P}(\mathbf{R})$ as required by the central limit theorem [144]. To discuss the statistical error in $\langle I \rangle$, we first have to introduce the variance of the function $f(\mathbf{R})$ given as the squared value of the standard deviation σ_f

$$\sigma_f^2 = \int (f(\mathbf{R}) - I)^2 d\mathbf{R},\tag{5.3}$$

from which the variance of the Monte Carlo estimate follows as

$$\sigma_{\langle I \rangle}^2 = \frac{\sigma_f^2}{L} \approx (L(L-1))^{-1} \sum_{i=1}^L (f(\mathbf{R}_i) - \langle I \rangle)^2\tag{5.4}$$

$$\approx L^{-2} \sum_{i=1}^L (f(\mathbf{R}_i)^2 - \langle I \rangle^2).\tag{5.5}$$

The standard deviation $\sigma_f L^{-\frac{1}{2}}$ can be interpreted as the estimated error bar for the Monte Carlo average $\langle I \rangle$, i.e. the value of I is in the interval

$$I - \sigma_f L^{-\frac{1}{2}} \leq \langle I \rangle \leq I + \sigma_f L^{-\frac{1}{2}}.\tag{5.6}$$

To obtain a confidence probability for this statement we have to resort to the central limit theorem: Providing a set of N independent stochastic sampled values $\{f_i\}$ with mean I and variance σ_f^2 obeying the same distribution function $F_L(f_i)$, the central limit theorem states that for large sampling sizes L the function $F_L(f_i)$ converges to a normal distribution

$$\lim_{L \rightarrow \infty} F_L(f_i) = \sqrt{L}(\sqrt{2\pi}\sigma_f)^{-1} \int_{-\infty}^{f_i} e^{-\frac{\sqrt{L}}{2\sigma_f}(t-I)^2} dt.\tag{5.7}$$

With eq. 5.7 at hand we obtain the confidence probability p_c of finding the value of $\langle I \rangle$ in the interval $[I - \sigma_f L^{-\frac{1}{2}}, I + \sigma_f L^{-\frac{1}{2}}]$ with the help of the error function

$$p_c = \sqrt{L}(\sqrt{2\pi}\sigma_f)^{-1} \int_{I - \sigma_f L^{-\frac{1}{2}}}^{I + \sigma_f L^{-\frac{1}{2}}} e^{-\frac{L(f_i - I)^2}{(2\sigma_f^2)}} df_i \quad (5.8)$$

$$= (\sqrt{2\pi})^{-1} \int_{-1}^{+1} e^{-\frac{1}{2}t^2} dt \quad (5.9)$$

$$= \text{erf}(2^{-\frac{1}{2}}) = 0.68268949. \quad (5.10)$$

This means that the probability of finding $\langle I \rangle$ in the expected interval is approx. 68%. Similarly we can introduce an expanded uncertainty $u_k = k\sigma_f L^{-\frac{1}{2}}$ with $k > 0$ to define the interval $I \pm u_k$ where we find from the error function the probability of $\langle I \rangle$ in $[I - k\sigma_f L^{-\frac{1}{2}}, I + k\sigma_f L^{-\frac{1}{2}}]$ as $[100 \cdot \text{erf}(k/\sqrt{2})]\%$. For example, the probability of being wrong by two error bars ($\pm u_2 = \pm 2\sigma_f L^{-\frac{1}{2}}$) is $p_c \approx 95\%$.

The independence condition for the samples is not a strict requirement. A central limit theorem is also ensured if the correlation between f_i and f_j goes sufficiently fast to zero for $|i - j| \rightarrow \infty$. Nevertheless, the samples in a standard Monte Carlo calculation are by far not independent, this means that the variance $\sigma_{\langle I \rangle}^2$ in eq. 5.5 does not give a valid uncertainty. A remedy for this defect is provided by the blocking technique which yields independent Gaussian stochastic variables by a recursive averaging of the samples. See Ref. [145] for a detailed description of the algorithm. Note that the effective error $\sigma_f L_{eff}^{-\frac{1}{2}}$ has to be expressed by the number of independent samples

$$L_{eff} = L\tau^{-1}, \quad (5.11)$$

where τ is the so called auto-correlation time, i.e. the average number of sampling steps to decorrelate the values of $f(\mathbf{R})$

$$\tau = 1 + \sum_{i=1}^{\infty} \frac{\langle f(\mathbf{R}_0) - \langle I \rangle \rangle \langle f(\mathbf{R}_i) - \langle I \rangle \rangle}{\sigma_f^2}. \quad (5.12)$$

From the error estimate we can see that the convergence of the result with $\sigma_f L^{-\frac{1}{2}}$ can be improved by reducing the variance σ_f^2 . This can be done by importance sampling which improves the sampling process significantly. If, for example, the distribution function $\mathcal{P}(\mathbf{R})$

gives large values in regions where $f(\mathbf{R})$ is low and vice versa, the Monte Carlo method will not yield reliable results for a finite sample size L because of the poor statistics. Introducing the normalized importance function $\mathcal{P}'(\mathbf{R})$ that exhibits a form similar to $f(\mathbf{R})$, we replace the score function $f(\mathbf{R})$ with

$$f'(\mathbf{R}) = \frac{f(\mathbf{R})}{\mathcal{P}'(\mathbf{R})}, \quad (5.13)$$

and sample the different configurations \mathbf{R} directly from $\mathcal{P}'(\mathbf{R})$. From eq. 5.2 we find the Monte Carlo estimate

$$\begin{aligned} \langle I \rangle' &= \int_{-\infty}^{\infty} \frac{f(\mathbf{R})}{\mathcal{P}'(\mathbf{R})} \mathcal{P}'(\mathbf{R}) d\mathbf{R} \\ &\approx L^{-1} \sum_{i=1}^L \frac{f(\mathbf{R}_i)}{\mathcal{P}'(\mathbf{R}_i)}. \end{aligned} \quad (5.14)$$

Obviously the average tends to I with $L \rightarrow \infty$. Considering the variance of eq. 5.5 we get

$$\frac{\sigma_f^2(\mathcal{P}')}{L} \approx L^{-2} \sum_{i=1}^L \left(\left[\frac{f(\mathbf{R}_i)}{\mathcal{P}'(\mathbf{R}_i)} \right]^2 - \langle I \rangle^2 \right), \quad (5.15)$$

i.e. we can significantly reduce the variance by a proper choice of the importance function close to $\mathcal{P}'(\mathbf{R}) \approx f(\mathbf{R})I^{-1}$.

5.1.1 Markov Chain Monte Carlo: The Metropolis Algorithm

In general we have to sample the set of configurations $\{\mathbf{R}\}$ from complex probability distributions $\mathcal{P}(\mathbf{R})$ with unknown normalization. The Metropolis rejection algorithm proposed in 1953 [146, 147] allows us to sample complex distributions with unknown normalization by the application of a Markov chain [148].

Let us first outline the Metropolis algorithm to sample a sequence of configurations $\{\mathbf{R}\}$. Starting from an initial configuration \mathbf{R}_i we determine a new position \mathbf{R}'_i with respect to a sampling probability $T(\mathbf{R}'_i \leftarrow \mathbf{R}_i)$. The new point \mathbf{R}'_i is accepted with probability

$$A(\mathbf{R}'_i \leftarrow \mathbf{R}_i) = \min \left[1, \frac{T(\mathbf{R}_i \leftarrow \mathbf{R}'_i) \mathcal{P}(\mathbf{R}'_i)}{T(\mathbf{R}'_i \leftarrow \mathbf{R}_i) \mathcal{P}(\mathbf{R}_i)} \right]. \quad (5.16)$$

If the trial configuration is accepted, \mathbf{R}'_i is the new point $\mathbf{R}_{i+1} = \mathbf{R}'_i$, otherwise \mathbf{R}_i is kept $\mathbf{R}_{i+1} = \mathbf{R}_i$. The transition probability for $\mathbf{R}'_i \leftarrow \mathbf{R}_i$ is given by

$$M(\mathbf{R}'_i | \mathbf{R}_i) = T(\mathbf{R}'_i \leftarrow \mathbf{R}_i) A(\mathbf{R}'_i \leftarrow \mathbf{R}_i), \quad (5.17)$$

where the constraints

$$\sum_j M(\mathbf{R}_j|\mathbf{R}_i) = 1, \quad M(\mathbf{R}_j|\mathbf{R}_i) \geq 0, \quad (5.18)$$

need to be satisfied. The stochastic evolution of the system is completely specified by the Markov matrix \mathbf{M} and a starting configuration \mathbf{R}_0 . Eq. 5.17 reflects the essential property of a Markov process, i.e. the probability of \mathbf{R}'_i emerging from \mathbf{R}_i only depends on the present state of the process. Let $\mathcal{P}_{n-1}(\mathbf{R})$ be the probability distribution after $(n-1)$ sampling steps, so we can express the evolution of $\mathcal{P}(\mathbf{R})$ as

$$\mathcal{P}_n(\mathbf{R}') = \sum_{\mathbf{R}} \mathcal{P}_{n-1}(\mathbf{R})M(\mathbf{R}'|\mathbf{R}), \quad (5.19)$$

which means that we can generate the whole sequence by subsequent application of matrix \mathbf{M} . Switching to a discrete representation we can write down the so called power law in the form of matrix-vector products

$$\mathcal{P}_n = \mathbf{M}\mathcal{P}_{n-1} = \mathbf{M}^n\mathcal{P}_0. \quad (5.20)$$

If the Markov process is ergodic, i.e. non-periodic and transitions between arbitrary states \mathbf{R} and \mathbf{R}' can be done in a finite number of steps, the distribution converges by eq. 5.20 to a unique equilibrium state. In the case that $\mathcal{P}(\mathbf{R})$ approaches equilibrium we can enforce the sufficient condition of detailed balance on the Markov Matrix

$$M(\mathbf{R}_i|\mathbf{R}_j)\mathcal{P}(\mathbf{R}_j) = M(\mathbf{R}_j|\mathbf{R}_i)\mathcal{P}(\mathbf{R}_i), \quad (5.21)$$

which ensures that we sample the desired distribution $\mathcal{P}(\mathbf{R})$, i.e. \mathcal{P} is the dominant right eigenvector of M . Note that the acceptance probability $A(\mathbf{R}'_i \leftarrow \mathbf{R}_i)$ is chosen in a way to fulfill eq. 5.21.

In order to elucidate the Metropolis algorithm we assume a large number of configurations at equilibrium state where the density of configurations in $d\mathbf{R}$ is given by $\rho(\mathbf{R})d\mathbf{R}$. Since eq. 5.21 holds, the transitions $d\mathbf{R} \rightarrow d\mathbf{R}'$ have to be balanced by transitions in the opposite direction

$$\begin{aligned} & A(\mathbf{R}'_i \leftarrow \mathbf{R}_i)T(\mathbf{R}'_i \leftarrow \mathbf{R}_i)\rho(\mathbf{R})d\mathbf{R}d\mathbf{R}' \\ & = A(\mathbf{R}_i \leftarrow \mathbf{R}'_i)T(\mathbf{R}_i \leftarrow \mathbf{R}'_i)\rho(\mathbf{R}')d\mathbf{R}'d\mathbf{R}. \end{aligned} \quad (5.22)$$

The equilibrium density is given by

$$\frac{\rho(\mathbf{R})}{\rho(\mathbf{R}')} = \frac{A(\mathbf{R}_i \leftarrow \mathbf{R}'_i)T(\mathbf{R}_i \leftarrow \mathbf{R}'_i)}{A(\mathbf{R}'_i \leftarrow \mathbf{R}_i)T(\mathbf{R}'_i \leftarrow \mathbf{R}_i)}, \quad (5.23)$$

resulting in

$$\frac{\rho(\mathbf{R})}{\rho(\mathbf{R}')} = \frac{\mathcal{P}(\mathbf{R})}{\mathcal{P}(\mathbf{R}')}, \quad (5.24)$$

using eq. 5.16. This means that the distribution of random walkers reflects the probability distribution $\mathcal{P}(\mathbf{R})$ we want to sample, generated without the knowledge of its normalization.

5.2 Variational Quantum Monte Carlo

Prior to the outline of the variational quantum Monte Carlo [149–151] method, we will reconsider the expectation value in terms of statistics. If we treat two electrons with space-spin coordinates x_1 and x_2 in the framework of classical mechanics the corresponding Coulomb potential in atomic units is easily found as $V_{12}(x_1, x_2) = |x_1 - x_2|^{-1}$. Switching to quantum mechanics the electron-pair configurations (x_1, x_2) are distributed according to $|\Psi(x_1, x_2)|^2$, i.e. the probability of finding electron 1 at x_1 and electron 2 at x_2 is proportional to the square of the wave function (Born's interpretation). So we have to sum up $V_{12}(x_1, x_2)$ from all (x_1, x_2) configurations weighted by their normed probabilities

$$V_{12} = \sum_{x_1, x_2} |x_1 - x_2|^{-1} \frac{|\Psi(x_1, x_2)|^2}{\sum_{x_1, x_2} |\Psi(x_1, x_2)|^2}, \quad (5.25)$$

where the analogy of the normalization to the grand partition function in statistical mechanics is obvious. Since the coordinates are continuous, we can rewrite eq. 5.25 as the familiar integral

$$V_{12} = \frac{\int \Psi(x_1, x_2)^* \hat{V}_{12} \Psi(x_1, x_2) dx_1 dx_2}{\int \Psi(x_1, x_2)^* \Psi(x_1, x_2) dx_1 dx_2}. \quad (5.26)$$

Introducing the probability distribution

$$\mathcal{P}(\mathbf{R}) = \frac{|\Psi(\mathbf{R})|^2}{\int |\Psi(\mathbf{R})|^2 d\mathbf{R}}, \quad (5.27)$$

describing the probability of definite electron configurations $\mathbf{R} = \{x_1, x_2, x_3, \dots, x_N\}$, the expectation value of an arbitrary operator \hat{O} is

$$\langle O \rangle = \int \frac{\hat{O}\Psi(\mathbf{R})}{\Psi(\mathbf{R})} \mathcal{P}(\mathbf{R}) d\mathbf{R}. \quad (5.28)$$

From the variational principle it follows that the expectation value for the ground-state energy E_0 formed with a trial wave function Ψ^T

$$E_0 = \frac{\int \Psi^{T*}(\mathbf{R}) \hat{H} \Psi^T(\mathbf{R}) d\mathbf{R}}{\int |\Psi^T(\mathbf{R})|^2 d\mathbf{R}}, \quad (5.29)$$

provides an upper bound to the exact energy $E_0^{exact} (\leq E_0)$. With the local energy E_{loc}

$$E_{loc}(\mathbf{R}) = \frac{\hat{H}\Psi^T(\mathbf{R})}{\Psi^T(\mathbf{R})}, \quad (5.30)$$

eq. 5.29 can be rewritten as

$$E_0 = \int E_{loc}(\mathbf{R}) \mathcal{P}(\mathbf{R}) d\mathbf{R}. \quad (5.31)$$

Using the Metropolis algorithm described in the previous section, L electron configurations \mathbf{R}_i are sampled from the probability density $\mathcal{P}(\mathbf{R})$ which give the energy expectation value as the average of the local energies

$$E_0 \approx L^{-1} \sum_{i=1}^L E_{loc}(\mathbf{R}_i). \quad (5.32)$$

With the use of sophisticated trial functions Ψ^T (see sec. 5.4) the VQMC method is able to yield accurate results comparable to established MO based correlation methods. Note that the trial function – aside from the standard constraints – must also be square integrable in order to provide proper statistics.

The success of the VQMC approach is based on the so-called zero-variance property. This means, that if the trial wave function approaches the exact ground-state function $\Psi^T \rightarrow \Psi_0^{ex}$, the local energy also approximates the exact ground-state energy $E_{loc}(\mathbf{R}) \rightarrow E_0^{ex}$, i.e. E_{loc} becomes independent of the given configuration \mathbf{R} with $\Psi^T \rightarrow \Psi_0^{ex}$. Thus the estimated variational energy converges more rapidly with respect to the number of Monte Carlo steps with increasing accuracy of the given trial function.

Even if projector QMC methods like fixed-node DQMC, which is described in section 5.3, are potentially more powerful than VQMC, the importance of the variational method

has to be stressed. These calculations are omnipresent in QMC applications since they are at least used to optimize the correlated wave function or density, respectively, i.e. the parameters of the correlation factors are optimized via correlated sampling [29, 152].

5.2.1 Importance Sampling by Langevin-type Fictitious Dynamics

In a simple VQMC algorithm the trial moves are directly sampled from a Gaussian distribution with standard deviation τ . As mentioned in the previous section, the sampling process can be improved by the use of importance sampling. A wide-spread method is motivated by the guide-function diffusion Monte Carlo method introducing fictitious Langevin-type dynamics. Then the proposal matrix $T(\mathbf{R}' \leftarrow \mathbf{R})$ in eq. 5.16 is given as

$$T(\mathbf{R}' \leftarrow \mathbf{R}) = (2\pi\tau)^{-3N/2} e^{-\frac{(\mathbf{R}' - \mathbf{R} - \mathbf{F}_q(\mathbf{R})\tau)^2}{2\tau}}, \quad (5.33)$$

with the quantum force

$$\mathbf{F}_q(\mathbf{R}) = \nabla \ln |\Psi^T(\mathbf{R})| = \Psi^T(\mathbf{R})^{-1} \nabla \Psi^T(\mathbf{R}). \quad (5.34)$$

where the Nabla operator ∇ acts on the electron coordinates. Sampling from eq. 5.33 corresponds to solving the Fokker-Planck equation

$$\frac{\partial}{\partial t} \mathcal{P}(\mathbf{R}) = \nabla (\nabla - \mathbf{F}_q(\mathbf{R})) \mathcal{P}(\mathbf{R}), \quad (5.35)$$

which describes the evolution of the probability distribution in time t . The Langevin equation to generate the trajectories according to eq. 5.35 is

$$\frac{\partial}{\partial t} \mathbf{R}(t) = \mathbf{F}_q(\mathbf{R}(t)) + \mathcal{G}, \quad (5.36)$$

with a random, Gaussian-distributed force \mathcal{G} . Integrating over a short time interval τ yields an equation suitable for generating trial configurations \mathbf{R}'

$$\mathbf{R}' = \mathbf{R} + \tau \mathbf{F}_q(\mathbf{R}) + \mathcal{G}_\tau. \quad (5.37)$$

The diffusion of random walkers is now partially directed by the quantum force into regions of large Ψ^T values but still exhibiting randomness (\mathcal{G}_τ) in order to be able to sample the whole configuration space.

Note that the discrete solution of eq. 5.36 introduces a time-step error for $\tau > 0$, which means that the results deteriorate with an increasing step size. The error may be estimated by extrapolation of results for different values of τ . Furthermore, other algorithms to ensure small time-step errors have been developed. At present the QUANTUMMC package [142] provides, besides the standard proposal matrix in eq. 5.33, the algorithm proposed by Umrigar, Nightingale and Runge [153] which exhibits small time-step errors.

5.3 Diffusion Quantum Monte Carlo

Apart from the VQMC method, where the integral in eq. 5.29 is sampled directly, there exist several potentially more powerful techniques that project out the ground state of the Hamiltonian (projector quantum Monte Carlo methods). In this brief section a short summary of the most popular of these methods, the fixed-node diffusion quantum Monte Carlo (FN-DQMC) algorithm, is given.

Starting from the energy-shifted time-dependent Schrödinger equation

$$\frac{\partial}{\partial t}\Psi(\mathbf{R}, t) = (\hat{H} - E_T)\Psi(\mathbf{R}, t), \quad (5.38)$$

with E_T as a constant energy offset and t as a real variable of imaginary time, we can transform this differential equation into integral form by means of the Green's function $G(\mathbf{R} \leftarrow \mathbf{R}', \tau)$ and the imaginary time variable τ

$$\Psi(\mathbf{R}, t + \tau) = \int G(\mathbf{R} \leftarrow \mathbf{R}', \tau)\Psi(\mathbf{R}', t)d\tau. \quad (5.39)$$

The Green's function has to obey the same time-dependence in eq. 5.38 like $\Psi(\mathbf{R}, t)$ and so results as

$$G(\mathbf{R} \leftarrow \mathbf{R}', \tau) = \langle \mathbf{R}' | e^{-\tau(\hat{H} - E_T)} | \mathbf{R} \rangle. \quad (5.40)$$

with $G(\mathbf{R} \leftarrow \mathbf{R}', 0) = \delta(\mathbf{R}' - \mathbf{R})$ as initial condition. Rewriting $G(\mathbf{R} \leftarrow \mathbf{R}', \tau)$ as well as the initial trial function $\Psi^{\text{init}}(\mathbf{R})$ in the spectral representation of the system

$$G(\mathbf{R} \leftarrow \mathbf{R}', \tau) = \sum_i |\Psi_i(\mathbf{R})\rangle e^{-\tau(E_i - E_T)} \langle \Psi_i(\mathbf{R}') |, \quad (5.41)$$

$$\Psi^{\text{init}}(\mathbf{R}) = \sum_i c_i \Psi_i(\mathbf{R}), \quad (5.42)$$

it is easily seen that the propagator $G(\mathbf{R} \leftarrow \mathbf{R}', \tau)$ projects out the lowest eigenstate Ψ_0 by evolving the system through imaginary time τ :

$$\begin{aligned} \lim_{\tau \rightarrow \infty} \int G(\mathbf{R} \leftarrow \mathbf{R}', \tau) \Psi^{\text{init}}(\mathbf{R}') d\mathbf{R}' &= \lim_{\tau \rightarrow \infty} \sum_i |\Psi_i(\mathbf{R})\rangle e^{-\tau(E_i - E_T)} \langle \Psi_i(\mathbf{R}') | \Psi^{\text{init}}(\mathbf{R}') \rangle \\ &= \lim_{\tau \rightarrow \infty} |\Psi_0(\mathbf{R})\rangle e^{-\tau(E_0 - E_T)} \langle \Psi_0(\mathbf{R}') | \Psi^{\text{init}}(\mathbf{R}') \rangle. \end{aligned} \quad (5.43)$$

Note that at least a non-zero overlap of the initial function with the exact ground state function must be provided ($c_0 > 0$). By adjusting E_T to E_0 the time-dependence of the ground-state term vanishes while the remaining overlaps are damped by $\exp[-\tau(E_i - E_T)]$. Note that $G(\mathbf{R} \leftarrow \mathbf{R}', \tau)$ is not a projector in the sense as discussed for \mathbf{P}_{occ} or \mathbf{P}_{virt} in section 2.5.2, i.e. it does not necessarily exactly project out the ground-state wave function. This would only hold if the energy shift E_T equals E_0 exactly, otherwise the other state functions are only damped compared to the ground-state function.

Since the fundamental properties of the propagator $G(\mathbf{R} \leftarrow \mathbf{R}', \tau)$ are at hand, we can briefly outline the DQMC algorithm for molecular systems. If we neglect the potential energy terms in the Hamiltonian (eq. 5.38), we obtain a master equation of a diffusion stochastic process $\partial_t \Psi(\mathbf{R}, t) = 1/2 \sum_i \nabla_i^2 \Psi(\mathbf{R}, t)$ where its Green's function is a Gaussian with variance τ

$$G(\mathbf{R} \leftarrow \mathbf{R}', \tau) = (2\pi\tau)^{3N/2} e^{-\frac{|\mathbf{R} - \mathbf{R}'|^2}{2\tau}}. \quad (5.44)$$

In the framework of stochastics the solution of the master equation $\Psi(\mathbf{R})$ describes the distribution of Brownian particles, so it can be represented by a discrete set of random walkers ($\Psi(\mathbf{R}) = \sum_i \delta(\mathbf{R} - \mathbf{R}_i)$).

To treat the complete Hamiltonian with potential terms we have to introduce the approximation of the Green's function introduced by Trotter and Suzuki [154, 155]. Within a short-time approximation ($\tau \rightarrow 0$) we obtain for eq. 5.40 the so-called primitive approximation

$$\begin{aligned} G(\mathbf{R} \leftarrow \mathbf{R}', \tau) &= \langle \mathbf{R}' | e^{-\tau(\hat{T} + \hat{V} - E_T)} | \mathbf{R} \rangle \\ &\approx (2\pi\tau)^{3N/2} e^{(2\tau)^{-1} |\mathbf{R} - \mathbf{R}'|^2} e^{-\frac{\tau}{2}(V(\mathbf{R}) - V(\mathbf{R}')) + \tau E_T}, \end{aligned} \quad (5.45)$$

where the error is cubic in the time-step size ($\sim \tau^3$). The exponential term containing the potential energy difference has the effect of a time-dependent renormalization of the diffusion process. While there exist several methods to consider this reweighting, our

QUANTUMMC code [142] uses a combination of assigning weights to the single random walkers and the birth-death algorithm as described in Ref. [153].

The algorithm described to this point is quite inefficient even for small systems, due to strong fluctuations in the reweighting term in eq. 5.45. To perform importance sampling [156], a trial or guiding function $|\Psi^T\rangle$ is used, which replaces Ψ with the product $f = \Psi\Psi^T$ in the foregoing equations. Inserting $f(\mathbf{R}, r)$ in eq. 5.38 we obtain

$$-\frac{\partial}{\partial t}f(\mathbf{R}, r) = -\frac{1}{2}\nabla^2 f(\mathbf{R}, r) + \nabla(\mathbf{F}_q(\mathbf{R})f(\mathbf{R}, r)) + (E_{\text{loc}}(\mathbf{R}) - E_T)f(\mathbf{R}, r), \quad (5.46)$$

with the quantum force $\mathbf{F}_q(\mathbf{R})^a$

$$\mathbf{F}_q(\mathbf{R}) = \nabla \ln |\Psi^T(\mathbf{R})| = \Psi^T(\mathbf{R})^{-1} \nabla \Psi^T(\mathbf{R}), \quad (5.47)$$

and the local energy E_{loc} as defined in eq. 5.30. Analogously to eq. 5.39 we obtain the integral equation

$$f(\mathbf{R}, t + \tau) = \int \tilde{G}(\mathbf{R} \leftarrow \mathbf{R}', \tau) f(\mathbf{R}', t) d\mathbf{R}', \quad (5.48)$$

where the corresponding Green's function $\tilde{G}(\mathbf{R} \leftarrow \mathbf{R}', \tau)$ results from comparison with eq. 5.39 as

$$\tilde{G}(\mathbf{R} \leftarrow \mathbf{R}', \tau) = \Psi^T(\mathbf{R}) G(\mathbf{R} \leftarrow \mathbf{R}', \tau) \Psi^T(\mathbf{R}')^{-1}. \quad (5.49)$$

In the short-time approximation we obtain

$$\tilde{G}(\mathbf{R} \leftarrow \mathbf{R}', \tau) = (2\pi\tau)^{3N/2} e^{(2\tau)^{-1}[\mathbf{R}-\mathbf{R}'-\tau\mathbf{F}_q(\mathbf{R}')]^2} e^{-\frac{\tau}{2}(E_{\text{loc}}(\mathbf{R})-E_{\text{loc}}(\mathbf{R}'))+\tau E_T}. \quad (5.50)$$

The branching term in eq. 5.50, i.e. the second exponential, now contains the local energy instead of the potential energy. If a good trial function is used, E_{loc} is close to the true ground-state energy and nearly constant, thus the population fluctuates only slightly within the sampling process. The energy is usually calculated with the mixed estimator E_m :

$$\begin{aligned} E_m &= \lim_{\tau \rightarrow \infty} \frac{\langle e^{-\frac{\tau}{2}\hat{H}}\Psi^T | \hat{H} | e^{-\frac{\tau}{2}\hat{H}}\Psi^T \rangle}{\langle e^{-\frac{\tau}{2}\hat{H}}\Psi^T | e^{-\frac{\tau}{2}\hat{H}}\Psi^T \rangle} \\ &= \lim_{\tau \rightarrow \infty} \frac{\langle e^{-\tau\hat{H}}\Psi^T | \hat{H} | \Psi^T \rangle}{\langle e^{-\tau\hat{H}}\Psi^T | \Psi^T \rangle} \\ &= \lim_{\tau \rightarrow \infty} \frac{\int f(\mathbf{R}, \tau) E_{\text{loc}}(\mathbf{R}) d\mathbf{R}}{\int f(\mathbf{R}, \tau) d\mathbf{R}} \\ &\approx L^{-1} \sum_l^L E_{\text{loc}}(\mathbf{R}_l). \end{aligned} \quad (5.51)$$

^{a)}In the literature the gradient in eq. 5.47 is often denoted as drift velocity \mathbf{v}_D .

Note that the single configurations are sampled from Ψ instead of $|\Psi|^2$, i.e. for fermionic systems we have to introduce a further approximation to take care of the antisymmetry of the wave function. Formally, the nodal regions are split with respect to the sign of the wave function, where an absorbing barrier is installed between two regions which can not be crossed by a random walker. The distribution can now be sampled by random walkers distributed into the different nodal pockets. If the locations of the nodes are exact, the calculation converges to the exact result [157]. The problem to locate the nodes can be approximately solved if importance sampling is applied^{b)}, i.e. one simply uses the nodes of the trial wave function Ψ^T . If a random walker crosses a node, the sign of Ψ^T changes. The trial position is rejected and the original configuration is sampled again.

The FN-DQMC methods have been proven to provide results of CCSD(T) quality [132] combined with a favorable scaling behavior as compared to standard MO-based correlation methods. The main error in the DQMC results from the approximated nodal structure of the system given by the trial function. Even if error cancellation in the calculation of energy differences occurs, overcoming the nodal problem is the main challenge in QMC method development. It has to be mentioned that some approaches exist that do not resort to the fixed-node approximation, but these are in general associated with a strongly increased computational effort.

5.4 Trial Wave Functions in Quantum Monte Carlo

In this section we will discuss the form of trial wave functions Ψ^T implemented in our QUANTUMMC package [142]. The quality of these functions is of course of central importance in VQMC calculations since the expectation value is directly sampled, but it is also important in FN-DQMC calculations since a proper nodal surface and sensible statistics have to be provided. The trial wave functions have the general Slater-Jastrow form [157]

$$\Psi_{\text{SJ}}^T(\mathbf{R}) = D_{\text{det}}^\alpha(\mathbf{R})D_{\text{det}}^\beta(\mathbf{R})e^{U(\mathbf{R})}, \quad (5.52)$$

where D_{det}^α and D_{det}^β are Slater determinants for α and β electrons built from molecular orbitals obtained from mean-field calculations. The separation of the determinant into

^{b)}Note that the force \mathbf{F}_q drives the random walker into regions of high importance, i.e. away from nodal regions. Thus the number of trial moves that would cross a node is a priori smaller if importance sampling is used.

D_{det}^α and D_{det}^β is easily done by choosing a permutation where the electrons are sorted corresponding to their spins. Since \mathbf{R} is only a dummy variable of integration and the operators are not affected by the permutation, we can replace the original, spin-dependent Slater determinant by the product of spin-independent determinants.

The exponential term e^U is the so called correlation or Jastrow factor containing the many-particle correlation factors. There are several approaches for U in the literature, but in the following section only the Boys-Handy factors of short-ranged type as proposed by Manten and Lüchow [135] will be described.

The evaluation of the trial function is the most demanding step in a QMC calculation scaling with $\mathcal{O}(N^3)$. Since the local energy is calculated with defined electron distributions, the potential energies are easily determined. However, the kinetic energy has to be calculated with $\Psi_{\text{SJ}}^{\text{T}}$ as $\Psi_{\text{SJ}}^{\text{T}-1}(\mathbf{R})\hat{T}\Psi_{\text{SJ}}^{\text{T}}(\mathbf{R})$. Considering the form of $\Psi_{\text{SJ}}^{\text{T}}$ in eq. 5.52 the kinetic energy results as

$$\begin{aligned}
 \Psi_{\text{SJ}}^{\text{T}-1}(\mathbf{R})\hat{T}\Psi_{\text{SJ}}^{\text{T}}(\mathbf{R}) &= -\frac{1}{2}\sum_i\frac{\nabla_i^2\Psi_{\text{SJ}}^{\text{T}}}{\Psi_{\text{SJ}}^{\text{T}}} \\
 &= -\frac{1}{2}\sum_i\frac{\nabla_i^2\Psi_{\text{SJ}}^{\text{T}}}{\Psi_{\text{SJ}}^{\text{T}}} + \left(\frac{\nabla_i\Psi_{\text{SJ}}^{\text{T}}}{\Psi_{\text{SJ}}^{\text{T}}}\right)^2 - \left(\frac{\nabla_i\Psi_{\text{SJ}}^{\text{T}}}{\Psi_{\text{SJ}}^{\text{T}}}\right)^2 \\
 &= -\frac{1}{2}\sum_i\nabla_i^2\ln|\Psi_{\text{SJ}}^{\text{T}}| - (\nabla_i\ln|\Psi_{\text{SJ}}^{\text{T}}|)^2 \\
 &= -\frac{1}{2}\sum_i\nabla_iU(\mathbf{R})(2D_{\text{det}}^{-1}\nabla_i(\mathbf{R})D_{\text{det}}(\mathbf{R}) + \nabla_iU(\mathbf{R})) \\
 &\quad -e^{-U(\mathbf{R})}\nabla_i^2e^{U(\mathbf{R})} - D_{\text{det}}^{-1}\nabla_i^2(\mathbf{R})D_{\text{det}}(\mathbf{R}), \tag{5.53}
 \end{aligned}$$

which requires the calculation of the gradient and Laplacian of the determinants and the correlation factors. The algorithms to compute the different ingredients to form eq. 5.53 will be discussed in the following sections.

It has to be mentioned that the construction of the trial wave function, i.e. how to improve on the IPM result, is a major task in QMC calculations. Even if for the examples presented in this work only a simple type of correlation factor is chosen, a good estimate to the true ground-state wave function is crucial in order to reduce the systematical and statistical errors. Their construction requires a relatively large amount of human time costs, i.e. QMC calculations are much less of a "black box" as compared to standard MO-based approaches such as coupled cluster calculations (see e.g. Ref. [158]).

5.4.1 The Slater Determinant

We will briefly discuss the evaluation of the local energy in the traditional VQMC method following Fahy et al. [149]. To calculate the gradient ($\Psi_{\text{SJ}}^{\text{T}^{-1}} \nabla_i \Psi_{\text{SJ}}^{\text{T}}$) and the Laplacian ($\Psi_{\text{SJ}}^{\text{T}^{-1}} \nabla^2 \Psi_{\text{SJ}}^{\text{T}}$) with respect to the single electron coordinates i , we first have to determine the values of basis functions $\chi_\mu(\mathbf{r})$ for all electron coordinates as well as the first and second derivatives.

Our program uses the same standard basis sets as common in ab initio program packages, where the radial parts of the basis functions are represented by cubic splines [105]. Instead of replacing s-type functions by Slater-type functions that exhibit a cusp at the nucleus, a fit to

$$f_{fit}(x) = ae^{-b|x|} + c \quad (5.54)$$

is applied for the region close to the nucleus (similar to Ref. [132]). As can be seen from the plot of the second derivatives in Fig. B.15, this fit clearly improves the behavior at the nucleus. Note that the user only has to mark the specific s-type function in the input file, so that this approach ensures an easy access to quantum Monte Carlo calculations. The calculation of the basis function values and their derivatives is formally an $\mathcal{O}(N^2)$ step, but – because of the exponential decay of the Gaussians – only a constant number of basis functions yield significant values for a single electron. To exploit this fact, a cut-off radius for each shell, for which the basis functions provide significant values, is determined with respect to the second derivative value, since these directly contribute to the kinetic energy [135]. Applying this screening with respect to a given threshold the computational effort in this step of the calculation scales linear with system size [159].

Following the traditional algorithm [149], the molecular orbital values are calculated according to the LCBF approach, where the AO-MO transformations scale cubically with system size

$$\phi_i(\mathbf{r}_j) = C_{\mu i} \chi_\mu(\mathbf{r}_j), \quad (5.55)$$

$$\nabla_j^{xyz} \phi_i(\mathbf{r}_j) = C_{\mu i} \nabla_j^{xyz} \chi_\mu(\mathbf{r}_j), \quad (5.56)$$

$$\nabla_j^2 \phi_i(\mathbf{r}_j) = C_{\mu i} \nabla_j^2 \chi_\mu(\mathbf{r}_j). \quad (5.57)$$

After the determinant matrix \mathbf{D}^c is built from $\{\phi_i\}$ to 5.58

$$\mathbf{D} = \begin{pmatrix} \phi_1(x_1) & \phi_2(x_1) & \cdots & \phi_N(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \cdots & \phi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(x_N) & \phi_2(x_N) & \cdots & \phi_N(x_N) \end{pmatrix}, \quad (5.58)$$

the inverse of its transposed form $\tilde{\mathbf{D}} = (\mathbf{D}^\dagger)^{-1}$ has to be calculated, i.e. the matrix of co-factors scaled by the inverse determinant. This is done within an $\mathcal{O}(N^3)$ LU decomposition which also provides the determinant values instantaneously by multiplying the diagonal elements of the factorized matrix [105]. The gradient and the Laplacian are evaluated by use of the Laplace expansion of the determinant in an $\mathcal{O}(N^2)$ scaling step

$$\Psi_{\text{SJ}}^{\text{T}-1} \nabla_i^{xyz} \Psi_{\text{SJ}}^{\text{T}} = \tilde{D}_{ji} \nabla_i^{xyz} \phi_j(\mathbf{r}_i), \quad (5.59)$$

$$\Psi_{\text{SJ}}^{\text{T}-1} \nabla^2 \Psi_{\text{SJ}}^{\text{T}} = \tilde{D}_{ji} \nabla^2 \phi_j(\mathbf{r}_i). \quad (5.60)$$

This expansion combined with Cramer's rule can also be used in an update procedure for $\tilde{\mathbf{D}}$ for single-electron moves [149, 157].

5.4.2 The Correlation Factor

Let us briefly reconsider the short-comings of the wave function in IPM calculations represented by a single Slater determinant. As shown in sec 2.6 the motion of two electrons with opposite spins is completely uncorrelated, i.e. the wave function does not provide a corresponding correlation cusp as shown in Fig. 2.3. Furthermore, since the basis functions are linear combinations of Gaussians, the mentioned electron-nuclear cusps are also missing. The true wave function, instead, has to satisfy a set of cusp conditions that were first derived by Kato [51]. Imagine two electrons that approach each other, so that the potential energy would diverge. This has to be in turn compensated by a corresponding divergence of the kinetic energy term in order to provide a finite total energy. The same holds for the potential energy arising from the electron-nuclear attraction. Consider the ground state energy of a hydrogen atom depending on the electron-nuclear distance r

$$-\frac{1}{2} \left(\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} \right) \Psi_0(r) - \frac{Z}{r} \Psi_0(r) = E \Psi_0(r). \quad (5.61)$$

^{c)}Note that the normalization factor $(N!)^{-\frac{1}{2}}$ is ignored since only the ratio of wave functions or expectation values are calculated.

Since the second derivatives are bounded [160], we have to demand

$$-\frac{1}{r} \left(\frac{d}{dr} + Z \right) \Psi_0(r) = \text{finite}, \quad (5.62)$$

in order to ensure the cancellation of divergences. With $\Psi_0(r) = e^{-\alpha r}$ we obtain

$$-\frac{1}{r} (-\alpha + Z) \Psi_0(r) = \text{finite}, \quad (5.63)$$

where it is easily seen that the choice $\alpha = Z$ satisfies the cusp conditions. From similar considerations one obtains for the electron-electron cusp

$$\lim_{r_{12} \rightarrow 0} \frac{\partial \Psi_{\text{SJ}}^{\text{T}}(\mathbf{R})}{\partial r_{12}} = \begin{cases} \frac{1}{2} \Psi_{\text{SJ}}^{\text{T}}(\mathbf{R}, r_{12} = 0) & \text{for } \sigma_1 \neq \sigma_2 \\ \frac{1}{4} \Psi_{\text{SJ}}^{\text{T}}(\mathbf{R}, r_{12} = 0) & \text{for } \sigma_1 = \sigma_2 \end{cases}, \quad (5.64)$$

where σ_i denotes the spin of electron i and r_{12} the inter-particle distance.

Even if there are different forms of correlation factors proposed by several authors, we will concentrate on the form suggested by Schmidt and Moskowitz [140]. They introduced a correlation factor originally proposed by Boys and Handy [141]

$$e^{U(\mathbf{R})} \rightarrow e^{\sum_{I,j} U_{I,ij}(\mathbf{R})} \\ U_{I,ij}(\mathbf{R}) = \sum_k \Delta_{mn} c_{kI} \left(\tilde{r}_{iI}^{m_{kI}} \tilde{r}_{jI}^{n_{kI}} + \tilde{r}_{iI}^{n_{kI}} \tilde{r}_{jI}^{m_{kI}} \right) \tilde{r}_{ij}^{o_{kI}}, \quad (5.65)$$

with the Padé-type scaled distances \tilde{r}

$$\tilde{r}_{ij} = \frac{d_I r_{ij}}{1 + d_I r_{ij}} \quad \text{and} \quad \tilde{r}_{iI} = \frac{b_I r_{iI}}{1 + b_I r_{iI}}, \quad (5.66)$$

for describing electron-electron and electron-nuclear correlations, respectively. Δ_{mn} ensures comparability [140] with the original work of Boys and Handy and c_{kI} are the coefficients of the linear combination of the N_I terms, the values of m , n and o are integers larger or equal to zero. A term with $(m = 0, n = 0, o \neq 0)$ for example describes the electron-electron correlation. Since the determinant does not depend on the inter-particle distance r_{12} , the cusp conditions have to be fulfilled by the correlation factor alone

$$\left. \frac{\partial \mathbf{U}(\mathbf{R})}{\partial r_{12}} \right|_{r_{12} \rightarrow 0} = \begin{cases} \frac{1}{2} & \text{for } \sigma_1 \neq \sigma_2 \\ \frac{1}{4} & \text{for } \sigma_1 = \sigma_2 \end{cases}. \quad (5.67)$$

Its effect is a reduction of the density close to the reference electron, but also results in a violation of the overall normalization which can be compensated by introducing pure

electron-nuclear terms ($(m \neq 0, n = 0)$ or $(m = 0, n \neq 0)$, $o = 0$). Note that a term like $(m = 1, n = 0, o = 0)$ could be included to improve the electron-nuclear cusp, but it becomes unnecessary by the fit of the basis functions discussed before.

In general, also higher-order correlation terms could be included [140], but we will only focus on a simplified version as proposed in [135] depicted in Tab. C.9. The scaled distances in eq. 5.66 provide a slow decay with $r \rightarrow \infty$, so that Manten and Lüchow introduced short-ranged distances

$$\tilde{r} = 1 - e^{-\alpha r}. \quad (5.68)$$

Note that eq. 5.68 provides the same functional form as the Padé-type function in eq. 5.66 for small values of r or r_{ij} , respectively. Since the correlation factor is only included to provide a proper behavior of the wave function for small particle distances, the strongly different decaying behavior of the function for $r \rightarrow \infty$ does not have a deteriorating effect compared to the Padé form in eq. 5.66. With the short-ranged distance at hand, the contributions of the correlation factor to the local energy can be screened with respect to the interparticle distances and a given threshold, so that a linear scaling behavior is achieved.

The traditional MO-based VQMC algorithm discussed so far has an $\mathcal{O}(N^3)$ scaling behavior for the calculation of the local energy E_{loc} . The scaling behavior results from the AO-MO transformations and the algebraic operations with the corresponding MO quantities (LU factorization etc.). In order to reduce the scaling Manten and Lüchow [135] use localized molecular orbitals [161–163] in combination with the described short-ranged correlation factor. However, the authors claimed to need a cut-off parameter in order to remove the orthogonalization tails and to obtain a linear scaling algorithm for the local energy E_{loc} .

In the following section the new N -PDM VQMC algorithm based on the one-electron density matrix is presented, which provides a route to an $\mathcal{O}(N)$ evaluation of E_{loc} without sacrificing accuracy.

5.5 N -Particle Density Matrix-based Variational Quantum Monte Carlo

Since we avoid the use of MOs in our N -PDM VQMC scheme by replacing them with the N -particle density ρ_N , we just need the discrete one-electron density matrix \mathbf{P} from a HF or KS-DFT calculation as well as the corresponding basis. As can be seen from eq. 2.22, the probability distribution $\mathcal{P}(\mathbf{R})$ and the local energy expression $E_{\text{loc}}(\mathbf{R})$ can be rewritten in terms of ρ_N :

$$\mathcal{P}(\mathbf{R}) = \frac{\rho_N(\mathbf{R})e^{2U(\mathbf{R})}}{\int \rho_N(\mathbf{R})e^{2U(\mathbf{R})}d\mathbf{R}}, \quad (5.69)$$

$$E_{\text{loc}}(\mathbf{R}) = \frac{\hat{H}\rho_N(\mathbf{R}; \mathbf{R}')e^{U(\mathbf{R})}e^{U(\mathbf{R}')}}{\rho_N(\mathbf{R}; \mathbf{R}')e^{U(\mathbf{R})}e^{U(\mathbf{R}')}}. \quad (5.70)$$

Again we use the standard convention from density matrix theory as explained in sec. 2.5, i.e. the Hamiltonian \hat{H} first acts on the set \mathbf{R}' , then we replace \mathbf{R}' with \mathbf{R} [43]. Because of the exponential form of the correlation factor no further modifications for the consideration of $e^{U(\mathbf{R})}$ have to be implemented. Thus we will focus on the determinant part of the calculation in the following text. Note that ρ_N is the product of α - and β -densities $\rho_N = \rho_N^\alpha \rho_N^\beta$, so that the following operations have to be done twice for ρ_N^α and ρ_N^β , respectively.

Initially we have to compute the basis functions values ($\chi_{i\mu} = \chi_\mu(\mathbf{r}_i)$) and their first ($\chi_{i\mu}^{u,1st} = \nabla_{u,i} \chi_\mu(\mathbf{r}_i)$, $u = x, y, z$) and second derivatives ($\chi_{i\mu}^{2nd} = \nabla_i^2 \chi_\mu(\mathbf{r}_i)$) for each electron using the same algorithm described in sec. 5.4.1. The determinant matrix \mathbf{D}_ρ ^{d)} according to eq. 2.35 is

$$\mathbf{D}_\rho = \begin{pmatrix} \rho_1(\mathbf{r}_1; \mathbf{r}'_1) & \rho_1(\mathbf{r}_1; \mathbf{r}'_2) & \cdots & \rho_1(\mathbf{r}_1; \mathbf{r}'_n) \\ \rho_1(\mathbf{r}_2; \mathbf{r}'_1) & \rho_1(\mathbf{r}_2; \mathbf{r}'_2) & \cdots & \rho_1(\mathbf{r}_2; \mathbf{r}'_n) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_1(\mathbf{r}_n; \mathbf{r}'_1) & \rho_1(\mathbf{r}_n; \mathbf{r}'_2) & \cdots & \rho_1(\mathbf{r}_n; \mathbf{r}'_n) \end{pmatrix}, \quad (5.71)$$

where r_i denote the single electron coordinates. This matrix is simply formed by two matrix multiplications

$$\rho_1(\mathbf{r}_i; \mathbf{r}'_j) = P_{\mu\nu} \chi_\mu(\mathbf{r}_i) \chi_\nu(\mathbf{r}'_j) \implies \mathbf{D}_\rho = \boldsymbol{\chi} \mathbf{P} \boldsymbol{\chi}^\dagger. \quad (5.72)$$

^{d)}Similar to eq. 5.58 the normalization factor $(N!)^{-1}$ is ignored.

Similarly we calculate

$$\begin{aligned} \mathbf{D}_\rho^{u,1st} &= \boldsymbol{\chi}^{u,1st} \mathbf{P} \boldsymbol{\chi}^\dagger \quad \text{with } u = x, y, z \\ \implies D_{\rho,ij}^{u,1st} &= \nabla_{u,i} \rho_1(\mathbf{r}_i; \mathbf{r}'_j), \end{aligned} \quad (5.73)$$

$$\begin{aligned} \mathbf{D}_\rho^{2nd} &= \boldsymbol{\chi}^{2nd} \mathbf{P} \boldsymbol{\chi}^\dagger \\ \implies D_{\rho,ij}^{2nd} &= \nabla_i^2 \rho_1(\mathbf{r}_i; \mathbf{r}'_j). \end{aligned} \quad (5.74)$$

Applying effective sparse algebra routines (see chapter 4) these six multiplies scale with $\mathcal{O}(N)$ as long as only matrices with a linear scaling number of significant elements occur. It has to be mentioned that a slightly different sparse matrix format is used compared to the RISSM implemented in the Q-Chem code. Since the matrices containing the basis function values are not quadratic, the compressed sparse row format (CSR) has been chosen in order to provide a much easier handling of $N \times M$ matrices [125]. Note that the arrangements of matrix multiplications in eq. 5.73 and eq. 5.74 are the most suitable for the CSR sparse format. The function to evaluate $\chi_{i\mu}$, $\chi_{i\mu}^{u,1st}$ and $\chi_{i\mu}^{2nd}$ directly scatters the results into sparse format, the following operations require only one matrix transposition.

The determination of the inverse $\tilde{\mathbf{D}}_\rho$ and the determinant could also be done with a LU decomposition. But in contrast to $\tilde{\mathbf{D}}$ which contains the MO values, we have a symmetric and positive definite^{e)} matrix allowing a Cholesky decomposition (see e.g. [105]) that in general is a factor of two faster than other factorizations [105]. Furthermore, an $\mathcal{O}(N)$ sparse matrix Cholesky decomposition routine [127] is by far easier to implement than a comparable sparse LU factorization [164, 165] because no pivoting has to be considered. Nevertheless, it has to be mentioned that small negative eigenvalues might occur resulting from the truncation of the matrices. Such trial configurations, which only have been observed at the equilibration stage in combination with loose sparse thresholds, are simply rejected.

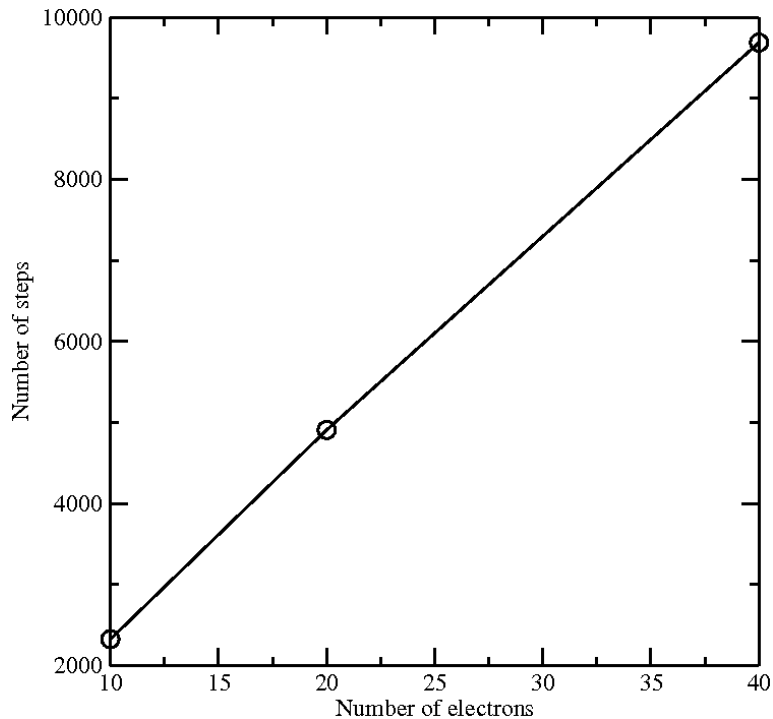
The gradient and the Laplacian are evaluated in a similar fashion to eqs. 5.59-5.60 by a sparse scalar product routine (overall $\mathcal{O}(N)$; $\mathcal{O}(1)$ for a single electron):

$$\rho_N(\mathbf{R}; \mathbf{R}')^{-1} \nabla_{u,i} \rho_N(\mathbf{R}; \mathbf{R}') = \tilde{D}_{\rho,ij} D_{\rho,ij}^{u,1st}, \quad (5.75)$$

$$\rho_N(\mathbf{R}; \mathbf{R}')^{-1} \nabla_i^2 \rho_N(\mathbf{R}; \mathbf{R}') = \tilde{D}_{\rho,ij} D_{\rho,ij}^{2nd}. \quad (5.76)$$

^{e)}Density matrices are of course positive semi-definite, but we treat a N_{el} configuration with a single zero diagonal element in \mathbf{D}_ρ as a $(N_{el} - 1)$ configuration. Thus the trial move is rejected before the Cholesky factorization takes place.

Figure 5.1: N -PDM VQMC calculations of a series of methane clusters (basis: cc-pVTZ). The number of steps to obtain a statistical error smaller than 10^{-3} is given for the calculation of a monomer, dimer and tetramer.



Since all quantities occurring in the described algorithm scale linear with system size (for systems exhibiting a significant HOMO-LUMO gap), an $\mathcal{O}(N)$ calculation of the local energy becomes possible, while rigorous error bounds are provided with respect to the given screening and sparse algebra thresholds. The scheme of the algorithm is depicted in Fig. B.16.

Altogether, this leads to an $\mathcal{O}(N^2)$ scaling behavior for calculations that provide the same statistical confidence (i.e. variance of the energy estimate). To elucidate this $\mathcal{O}(N^2)$ behavior one has to consider a molecular system with a significant HOMO-LUMO gap. Since the electrons are locally distributed, one can split the local energy into single electron contributions that are approximately independent [157]. Since each contribution also has an average variance of σ_e^2 , the total variance is proportional to the number of electrons N . Combined with the $\mathcal{O}(N)$ evaluation of the trial density we see that the N -PDM VQMC algorithm scales quadratically with system size.

As an example, the ground-state energies of different methane clusters have been calculated. For all calculations the same correlation factors are used, which have been obtained by variance minimization of the methane monomer. The initial random walkers for the dimer and tetramer systems are generated from a N -PDM VQMC calculation of the monomer by duplication. Since the different molecules of the clusters are well separated by at least 500 Å, the intermolecular interactions are negligible. Thus the overall variance can be written as a sum of the single variances σ_i^2 of each molecule

$$\sigma^2 = \sum_i^{N_{\text{mol}}} \sigma_i^2. \quad (5.77)$$

Because of the equality of the subsystems we can write

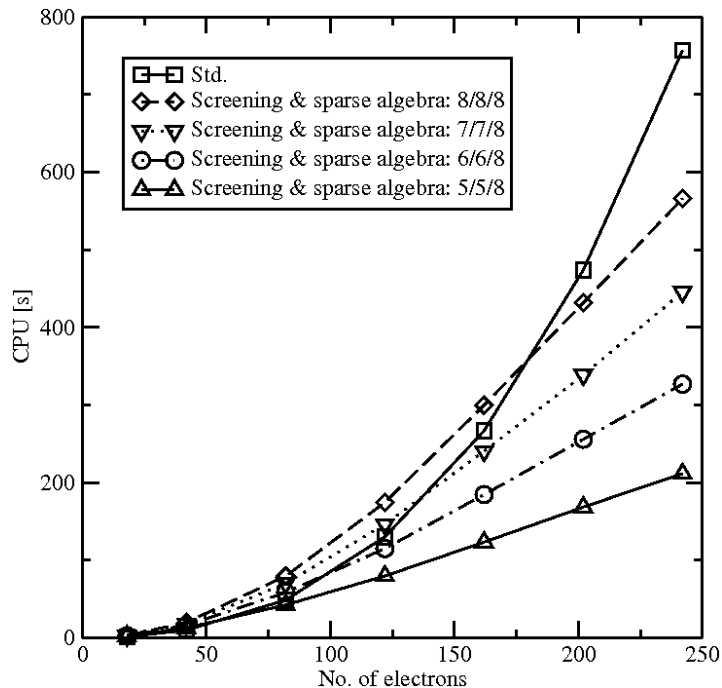
$$\sigma^2 = N_{\text{mol}} \sigma_{\text{monomer}}^2, \quad (5.78)$$

i.e. the overall variance is proportional to the number of methane molecules. Therefore, as shown in Fig. 5.1, the number of steps to obtain a statistical error smaller than 10^{-3} grows linearly with the number of electrons.

5.5.1 Illustrative Calculations

The scaling behavior of the new N -PDM VQMC method is shown for a series of linear alkanes (C_2H_6 , C_5H_{12} , $\text{C}_{10}\text{H}_{22}$, $\text{C}_{15}\text{H}_{32}$, $\text{C}_{20}\text{H}_{42}$, $\text{C}_{25}\text{H}_{52}$, $\text{C}_{30}\text{H}_{62}$). The determinant part of the correlated wave function is constructed from the molecular orbitals of a Hartree-Fock calculation with a cc-pVTZ basis [166]. The largest system ($\text{C}_{30}\text{H}_{62}$) contains 242 electrons and 1980 cartesian basis functions. Since these calculations should only prove the linear scaling behavior of the new method, the simple short-ranged correlation factor EJ_5 [135] given in Tab. C.9 is used. It has to be mentioned that an optimized class for the EJ_5 set of terms is implemented in QUANTUMMC [142] which ensures a rapid computation of the correlation factor contributions to the energy and the wave function. Thus the evaluation of the determinant \mathbf{D}_ρ is the by far dominant step in the calculation. The correlation factors are determined in a similar way to Manten and Lüchow [135], who set the values of α in eq. 5.68 for the higher-order terms 2-5 (see Tab. C.9) equal to 3.0. The linear coefficients as well as the exponential coefficient in term 1 are optimized within a correlated sampling algorithm according to Umrigar et al. [152].

Figure 5.2: N -PDM VQMC calculations of a series of linear alkanes (basis: cc-pVTZ). CPU times for for 1000 sampling steps of a single random walker. For an explanation of the integer code ($x/y/z$) see the text.



All computations have been performed with the QUANTUMMC program [142], which has been developed in the present PhD thesis, on a single Intel Xeon EM64T 3.0 GHz processor. After an equilibration phase of 10000 steps, the CPU times for a single random walker and 1000 sampling steps were measured and are shown in Fig. 5.2. The sampling algorithm proposed by Umrigar et al. [153] has been used where all electrons are moved simultaneously. The combination of numbers $x/y/z$ in the legend of Fig. 5.2 denotes the different thresholds for basis function screening ($\text{thr}_{\text{BF}} = 10^{-x}$), for the compression of the discrete one-electron density matrix \mathbf{P} ($\text{thr}_{\mathbf{P}} = 10^{-y}$), and a general sparse matrix compression threshold ($\text{thr}_{\text{SA}} = 10^{-z}$). The need for a special truncation criterion for the discrete density matrix arises from the structure of the quantities occurring in the density-based equations. The sparse determinant matrix \mathbf{D}_{ρ} has in general small off-diagonal elements compared to the diagonal elements which yields an inverse containing elements of large absolute value. Since the expectation values are calculated according to eq. 5.75 and eq. 5.76, even smaller values of $\mathbf{D}_{\rho}^{u,1st}$ and \mathbf{D}_{ρ}^{2nd} have to be considered in product with $\tilde{\mathbf{D}}_{\rho}$ (see

Table 5.1: Scaling behavior $\mathcal{O}(N^x)$ of the number of significant elements in the one-electron density matrix \mathbf{P} and the total CPU time for the different calculations. For a comment on the accuracy of the scaling behavior of \mathbf{P} and an explanation of the integer code (x/y/z) see the text.

	5/5/8		6/6/8		7/7/8		8/8/8	
	\mathbf{P}	Tot. time	\mathbf{P}	Tot. time	\mathbf{P}	Tot. time	\mathbf{P}	Tot. time
$\text{C}_2\text{H}_6\text{-C}_5\text{H}_{12}$	1.7	2.1	1.8	2.2	1.9	2.2	1.9	2.2
$\text{C}_5\text{H}_{12}\text{-C}_{10}\text{H}_{22}$	1.6	1.8	1.7	1.9	1.9	2.0	2.0	2.1
$\text{C}_{10}\text{H}_{22}\text{-C}_{15}\text{H}_{32}$	1.3	1.6	1.4	1.7	1.6	1.8	1.8	2.0
$\text{C}_{15}\text{H}_{32}\text{-C}_{20}\text{H}_{42}$	1.2	1.5	1.4	1.7	1.5	1.7	1.7	1.9
$\text{C}_{20}\text{H}_{42}\text{-C}_{25}\text{H}_{52}$	1.1	1.4	1.2	1.5	1.3	1.5	1.5	1.7
$\text{C}_{25}\text{H}_{52}\text{-C}_{30}\text{H}_{62}$	1.1	1.3	1.2	1.4	1.2	1.5	1.4	1.5

discussion in section 4.3). In the construction of these intermediate matrices like $\mathbf{D}_\rho^{u,1st}$ one has to choose a tighter compression threshold thr_{SA} while a more crude truncation of \mathbf{P} with respect to $\text{thr}_{\mathbf{P}} > \text{thr}_{\text{SA}}$ does not result in a larger deterioration. Note that the truncation of the density matrix \mathbf{P} results in an electron loss

$$\text{Tr}[\mathbf{P}_{\text{trunc}}\mathbf{S}] \neq N, \quad (5.79)$$

thus \mathbf{P} has been renormalized to N with the McWeeny purification transformation in eq. 2.51. As one can see in Fig. B.18 in the appendix, the choice of a less tight threshold for \mathbf{P} has a large impact on the computation time. The computation with $\text{thr}_{\text{SA}} = \text{thr}_{\mathbf{P}} = 10^{-8}$ is approx. 3 times slower than the computation with $\text{thr}_{\text{SA}} = 10^{-8}$, $\text{thr}_{\mathbf{P}} = 10^{-5}$. Since the multiplies take approximately 90% of the total CPU time, a less tight density compression threshold $\text{thr}_{\mathbf{P}}$ significantly reduces the total timings presented in Fig. 5.2. It has to be noted that the time for the Cholesky factorization as well as the contraction of e.g. eq. 5.75 are nearly negligible. In Tab. 5.1 the scaling behavior of the number of significant elements of \mathbf{P} and of the total CPU time for the thousand sampling steps are compared. It has to be mentioned that the number of significant elements of \mathbf{P} are taken from the standard output which prints the sparsity for different thresholds in percent. Since only one digit is given, the scaling behavior of \mathbf{P} in Tab. 5.1 shows an increasing inaccuracy with

Table 5.2: Error estimation via correlated sampling for different density compression thresholds (first value) and the general sparse algebra threshold thr_{SA} (second value). The screening threshold is set equal to the density compression threshold ($\text{thr}_{\text{BF}} = \text{thr}_{\text{P}}$). The value in brackets gives the uncertainty in the last given digit.

thr_{P}	thr_{SA}				
	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-10}
10^{-5}	—			0.0003(0)	0.0003(0)
10^{-6}		0.01(0)		0.00009(0)	0.00008(0)
10^{-7}			0.0003(0)	0.00002(1)	0.000007(0)
10^{-8}				0.000004(0)	0.0000005(0)

system size. The scaling behavior for e.g. $\text{C}_{25}\text{H}_{52}\text{-C}_{30}\text{H}_{62}$ exhibits an uncertainty of ± 0.1 . Nevertheless, one can clearly see that the overall scaling behavior is strongly coupled to the scaling behavior of the one-electron density matrix \mathbf{P} , even though the sparse algebra threshold thr_{SA} is chosen significantly more tight (5/5/8 or 6/6/8).

The evaluation of the basis functions only requires a small part of the total CPU time as is seen in Fig. B.17 in the appendix. Since the Gauss functions exhibit an exponential decay, even a tight threshold of $\text{thr}_{\text{BF}} = 10^{-8}$ accelerates the algorithm significantly. It has to be mentioned that small fluctuations in the CPU times for the smaller systems occur in Fig. B.17 which can be easily traced back to inaccuracies in the single measured times. Since the plotted values are the sum of 1000 sampling steps, a single evaluation took at most 0.02 seconds for $\text{C}_{30}\text{H}_{62}$ with (6/6/8).

In order to analyze the influence of the matrix truncation and the basis function screening the energy difference compared to a standard $\mathcal{O}(N^3)$ computation has been estimated via correlated sampling [29], the results are presented in Tab. 5.2. It is clear from intuition that the matrix truncation has the largest deteriorating impact on the results because of its omnipresence in the algorithm. So it was even not possible to generate a Markov chain with a sparse algebra threshold of $\text{thr}_{\text{SA}} = 10^{-5}$ since actually all trial moves have been rejected, i.e. nearly all elements of \mathbf{D}_ρ vanished. From Tab. 5.2 one can see that the thresholds 6/6/8 lead to an error in the order of 0.1 mHartree, a tight sparse threshold of 10^{-10}

does only result in an improved accuracy when $\text{thr}_{\mathbf{P}}$ and thr_{BF} are lowered simultaneously. The necessity of a tight sparse threshold thr_{SA} is clearly seen from the first three columns with $\text{thr}_{\mathbf{P}} = \text{thr}_{\text{BF}} = \text{thr}_{\text{SA}}$.

Finally, one can state that a N -PDM VQMC calculation for linear alkanes with 6/6/8 yields results of 0.1 mHartree accuracy combined with an early onset of the linear scaling behavior for a fixed sample size. Compared to a localized molecular orbital approach [135] the density-based algorithm does not only provide rigorous error bounds but also allows the application of the efficient sparse Cholesky decomposition [127]. As we will see in section 5.6.1, a sparse LU factorization [164] is less efficient and exhibits an overhead of two orders of magnitude for the presented examples (see Fig. 5.4 on page 126).

5.6 N -Particle Density Matrix-based Diffusion Quantum Monte Carlo

In order to derive analogous equations for the popular FN-DQMC method in terms of the N -particle density ρ_N , we have to reconsider the treatment of the fermion sign problem in the conventional algorithm. As it has been mentioned in sec. 5.3, the use of the fixed-node approximation is crucial in order to prevent an exponential decaying signal-to-noise ratio [157]. The importance-sampled Green's function $G(\mathbf{R} \leftarrow \mathbf{R}')$ in eq. 5.49 fulfills the detailed balance condition

$$\tilde{G}(\mathbf{R} \leftarrow \mathbf{R}') \left(\Psi_{\text{SJ}}^{\text{T}}(\mathbf{R}') \right)^2 = \tilde{G}(\mathbf{R}' \leftarrow \mathbf{R}) \left(\Psi_{\text{SJ}}^{\text{T}}(\mathbf{R}) \right)^2. \quad (5.80)$$

Thus the Metropolis acceptance probability $A(\mathbf{R} \leftarrow \mathbf{R}')$ in eq. 5.16 is given as

$$\begin{aligned} A(\mathbf{R} \leftarrow \mathbf{R}') &= \min \left[1, \frac{\tilde{G}(\mathbf{R}' \leftarrow \mathbf{R}) \Psi_{\text{SJ}}^{\text{T}}(\mathbf{R})^2}{\tilde{G}(\mathbf{R} \leftarrow \mathbf{R}') \Psi_{\text{SJ}}^{\text{T}}(\mathbf{R}')^2} \right] \\ &= \min \left[1, \frac{e^{-(2\tau)^{-1}(\mathbf{R}' - \mathbf{R} - \tau \mathbf{F}_q(\mathbf{R}))^2}}{e^{-(2\tau)^{-1}(\mathbf{R} - \mathbf{R}' - \tau \mathbf{F}_q(\mathbf{R}'))^2}} \left(\frac{\Psi_{\text{SJ}}^{\text{T}}(\mathbf{R})}{\Psi_{\text{SJ}}^{\text{T}}(\mathbf{R}')} \right)^2 \right], \end{aligned} \quad (5.81)$$

where the second exponential term of \tilde{G} in eq. 5.50, which contains the local energies, vanishes because of its symmetry with respect to the exchange of \mathbf{R} and \mathbf{R}'

$$e^{-\frac{\tau}{2}(E_{\text{loc}}(\mathbf{R}) - E_{\text{loc}}(\mathbf{R}')) + \tau E_T} = e^{-\frac{\tau}{2}(E_{\text{loc}}(\mathbf{R}') - E_{\text{loc}}(\mathbf{R})) + \tau E_T}. \quad (5.82)$$

In the standard approach, the ratio of the wave function values $\Psi_{\text{SJ}}^{\text{T}}(\mathbf{R})$ and $\Psi_{\text{SJ}}^{\text{T}}(\mathbf{R}')$ is used to impose the fixed-node approximation. If \mathbf{R} and \mathbf{R}' are in different nodal pockets with opposite sign, the ratio is negative and the trial position \mathbf{R} is discarded. It is obvious that we could easily replace the squared wave function values in eq. 5.81 with the density values

$$\left(\frac{\Psi_{\text{SJ}}^{\text{T}}(\mathbf{R})}{\Psi_{\text{SJ}}^{\text{T}}(\mathbf{R}')}\right)^2 = \frac{\rho_N(\mathbf{R})}{\rho_N(\mathbf{R}')}, \quad (5.83)$$

but this would of course destroy the information on the sign of the nodal pockets.

In order to provide a remedy to this problem, we reformulate the different steps of the FN-DQMC algorithm in terms of off-diagonal elements of the density $\rho_N(\mathbf{R}; \mathbf{R}')$. Note that this choice seems to be natural if we reconsider the Green's function as a thermal density as it is done in path integral quantum Monte Carlo methods (PIQMC) [167, 168]. By translating this idea to the FN-DQMC algorithm we will be able to obtain the necessary nodal information from the sign of the trial density $\rho_N(\mathbf{R}; \mathbf{R}')$ since

$$\text{sign}(\rho_N(\mathbf{R}; \mathbf{R}')) = \text{sign}(\Psi_{\text{SJ}}^{\text{T}}(\mathbf{R})) \cdot \text{sign}(\Psi_{\text{SJ}}^{\text{T}*}(\mathbf{R}')). \quad (5.84)$$

The algorithm as it is implemented in our QUANTUMMC package [142] is outlined in the following (see also scheme in Fig. B.20).

Starting from an initial configuration \mathbf{R}_α , the local energy, the gradient, the Laplacian and the diagonal element $\rho_N(\mathbf{R}_\alpha; \mathbf{R}_\alpha)$ are calculated as described for the density-based VQMC algorithm in sec. 5.5. A new trial configuration is proposed as

$$\mathbf{R}_{\alpha+1} = \mathbf{R}_\alpha + \tau \mathbf{F}_q(\mathbf{R}_\alpha) + \mathcal{G}_\tau, \quad (5.85)$$

but instead of proceeding analogously to the VQMC algorithm, we determine the necessary expectation values with off-diagonal elements of the density matrix

$$\langle \hat{O}(\mathbf{R}_{\alpha+1}) \rangle = \rho_N(\mathbf{R}_{\alpha+1}; \mathbf{R}_\alpha)^{-1} \hat{O} \rho_N(\mathbf{R}_{\alpha+1}; \mathbf{R}_\alpha), \quad (5.86)$$

where the operator acts on $\mathbf{R}_{\alpha+1}$ only. After the evaluation of the basis function values

and their derivatives for $\mathbf{R}_{\alpha+1}$, the following matrices are built:

$$\begin{aligned} \mathbf{D}_\rho(\mathbf{R}_{\alpha+1}; \mathbf{R}_\alpha) &= \boldsymbol{\chi}(\mathbf{R}_{\alpha+1})\mathbf{P}(\boldsymbol{\chi}(\mathbf{R}_\alpha))^\dagger \\ \implies \rho_1(\mathbf{r}_i^{\alpha+1}; \mathbf{r}_j^\alpha) &= P_{\mu\nu}\chi_\mu(\mathbf{r}_i^{\alpha+1})\chi_\nu(\mathbf{r}_j^\alpha), \end{aligned} \quad (5.87)$$

$$\begin{aligned} \mathbf{D}_\rho^{u,1st}(\mathbf{R}_{\alpha+1}; \mathbf{R}_\alpha) &= \boldsymbol{\chi}^{u,1st}(\mathbf{R}_{\alpha+1})\mathbf{P}(\boldsymbol{\chi}(\mathbf{R}_\alpha))^\dagger \quad \text{with } u = x, y, z \\ \implies D_{\rho,ij}^{u,1st}(\mathbf{R}_{\alpha+1}; \mathbf{R}_\alpha) &= \nabla_{u,i}\rho_1(\mathbf{r}_i^{\alpha+1}; \mathbf{r}_j^\alpha), \end{aligned} \quad (5.88)$$

$$\begin{aligned} \mathbf{D}_\rho^{2nd}(\mathbf{R}_{\alpha+1}; \mathbf{R}_\alpha) &= \boldsymbol{\chi}^{2nd}(\mathbf{R}_{\alpha+1})\mathbf{P}(\boldsymbol{\chi}(\mathbf{R}_\alpha))^\dagger \\ \implies D_{\rho,ij}^{2nd}(\mathbf{R}_{\alpha+1}; \mathbf{R}_\alpha) &= \nabla_i^2\rho_1(\mathbf{r}_i^{\alpha+1}; \mathbf{r}_j^\alpha). \end{aligned} \quad (5.89)$$

The expectation values are obtained by a Laplace expansion of the corresponding determinants as in eq. 5.75 and eq. 5.76. In contrast to the determinant matrix \mathbf{D}_ρ in eq. 5.71 we have to deal with a non-symmetric matrix $\mathbf{D}_\rho(\mathbf{R}_{\alpha+1}; \mathbf{R}_\alpha)$. Thus we cannot use the Cholesky decomposition and have to resort to the LU factorization as it is used in the conventional or the LMO-based algorithm [135, 149]. As it has been mentioned before, the treatment of sparse matrices in a linear scaling LU factorization is by far more complicated than in the Cholesky decomposition since pivoting has to be used. However, the UMF-PACK library [165] provides a linear scaling routine which also supports the compressed sparse row format (CSR) as used in our QUANTUMMC package. Since the number of significant elements in all occurring matrices exhibit a linear scaling behavior with system size, an overall $\mathcal{O}(N)$ scaling for the local energy is possible.

The last problem that has to be solved is the calculation of the ratio $\Psi(\mathbf{R}_{\alpha+1})/\Psi(\mathbf{R}_\alpha)$. The determination of the nodes is easily done by tracing the sign of the density $\rho_N(\mathbf{R}_{\alpha+1}; \mathbf{R}_\alpha)$, i.e. its value is negative if the random walker moves into a nodal pocket of opposite sign. The ratio is simply calculated as

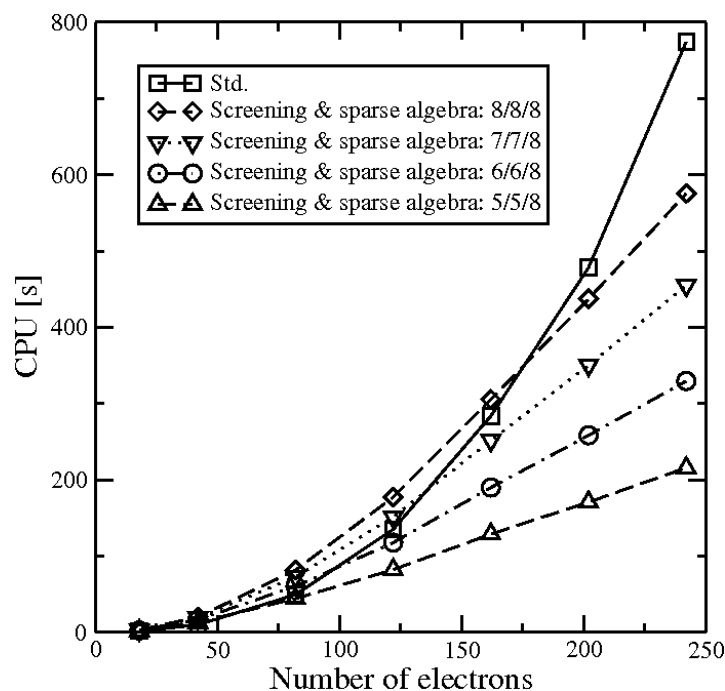
$$\begin{aligned} \frac{\rho_N(\mathbf{R}_{\alpha+1}; \mathbf{R}_\alpha)}{\rho_N(\mathbf{R}_\alpha; \mathbf{R}_{\alpha-1})} \left(\frac{\Psi(\mathbf{R}_\alpha)}{\Psi(\mathbf{R}_{\alpha-1})} \right)^{-1} &= \frac{\Psi(\mathbf{R}_{\alpha+1})\Psi(\mathbf{R}_\alpha)}{\Psi(\mathbf{R}_\alpha)\Psi(\mathbf{R}_{\alpha-1})} \frac{\Psi(\mathbf{R}_{\alpha-1})}{\Psi(\mathbf{R}_\alpha)} \\ &= \frac{\Psi(\mathbf{R}_{\alpha+1})}{\Psi(\mathbf{R}_\alpha)}, \end{aligned} \quad (5.90)$$

with the initial ratio

$$\frac{\Psi(\mathbf{R}_1)}{\Psi(\mathbf{R}_0)} = \frac{\rho_N(\mathbf{R}_1; \mathbf{R}_0)}{\rho_N(\mathbf{R}_0; \mathbf{R}_0)}. \quad (5.91)$$

When the matrix $\mathbf{P}\boldsymbol{\chi}^\dagger$ is kept in memory, the effort to compute a single time step in the density matrix-based FN-DQMC algorithm is similar to the presented VQMC algorithm.

Figure 5.3: N -PDM FN-DQMC calculations of a series of linear alkanes (basis: cc-pVTZ). CPU times for for 1000 sampling steps of a single random walker. The standard $\mathcal{O}(N^3)$ routine for the LU factorization [105] is used. For an explanation of the integer code (x/y/z) see the text on page 119.

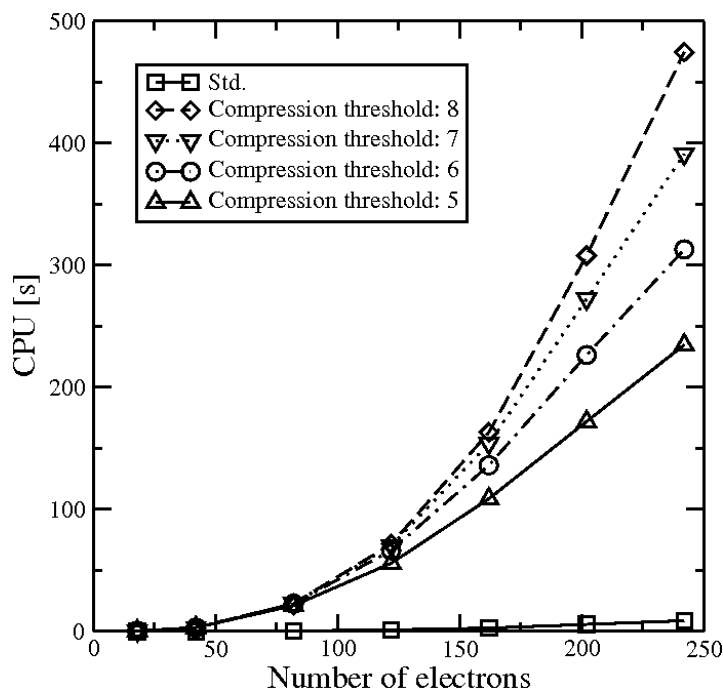


5.6.1 Illustrative Calculations

As for the N -PDM VQMC algorithm, the scaling behavior is shown for the example of linear alkanes. Analogously to the calculations presented in sec. 5.5.1, we used the one-particle density matrices obtained by HF/cc-pVTZ calculations and the same correlation factors are used as in the N -PDM VQMC calculations. The N -PDM FN-DQMC calculations have been performed on a single Intel Xeon EM64T 3.0 GHz processor with our QUANTUMMC program [142] using the sampling algorithm of Umrigar et al. [153].

As it has been shown in the preceding section, the linear scaling computation of the local energy in FN-DQMC is possible by exploiting the sparsity of the occurring matrices by efficient sparse algebra routines. Unfortunately, the LU factorization of sparse matrices is much more involved because of the pivoting [164]. Since such a routine has not been

Figure 5.4: N -PDM FN-DQMC calculations of a series of linear alkanes (basis: cc-pVTZ). CPU times needed for the determination of the inverse $\tilde{\mathbf{D}}_\rho^\dagger$ and the determinant value by LU factorization. The CPU is measured for a single random walker and 1000 sampling steps. The compression threshold for the determinant matrix $\tilde{\mathbf{D}}_\rho(\mathbf{R}; \mathbf{R}')$ is given by the integer x in the legend (10^{-x}).



developed in this work, we first circumvented the problem by using the standard $\mathcal{O}(N^3)$ algorithm [105]. From the N -PDM VQMC calculations in section 5.5.1 we know that the prefactor of the Cholesky decomposition is small, so we expect the same behavior for the $\mathcal{O}(N^3)$ LU factorization. Therefore, the dense determinant matrix $\tilde{\mathbf{D}}_\rho(\mathbf{R}; \mathbf{R}')$ is reconstructed from the sparse representation and treated within the $\mathcal{O}(N^3)$ routine to form the determinant and inverse $(\tilde{\mathbf{D}}_\rho(\mathbf{R}; \mathbf{R}'))^\dagger$, which is then compressed to sparse storage format. The results for this approach are depicted in Fig. 5.3. Compared to the N -PDM VQMC results in Fig. 5.2 on page 119 one can see that the N -PDM VQMC and N -PDM FN-DQMC algorithms took approximately the same amount of CPU time as one would also expect from the equations in section 5.5 and 5.6.

Because of the small prefactor of the $\mathcal{O}(N^3)$ LU routines we obtained an overall linear scaling behavior for the computation of the local energy E_{loc} (see also Fig. 5.4). For larger systems this step would of course dominate the computation time, thus we also tested the routines of the UMFPACK library [165]. As be seen from the results shown in Fig. 5.4 and Fig. B.21, the LU routines are extremely inefficient for the calculated systems. The prefactor of the sparse routine is at least 25 for $\text{C}_{30}\text{H}_{62}$ if the determinant matrix is re-compressed with respect to $\text{thr} = 10^{-5}$ before the LU factorization. Note that the sparse algebra threshold is 10^{-8} for each calculation, thus the scaling behavior of the LU factorization clearly dominates the overall scaling as shown in Fig. B.21. Even if the determinant matrix is re-compressed before the LU factorization is applied, the sparse algebra calculations are only slightly faster than the $\mathcal{O}(N^3)$ calculation (see Fig. B.22).

The later onset of the sparse LU factorization may explain the claim of Manten and Lüchow [135] to need a cut-off parameter for the construction of the localized molecular orbitals. Even if the LMO determinant matrix exhibits a more sparse structure than the density determinant, a further truncation could have been necessary in order to achieve a linear scaling behavior for the treated system size. We also experience this problem of a late onset of the $\mathcal{O}(N)$ behavior for the sparse LU algorithm. However, since the prefactor of the routine to construct the inverse and the determinant is small compared to the matrix multiplications, we have to consider the application of sparse LU routines only for larger system sizes than those presented in this work. It has to be mentioned further that the UMFPACK library treats sparse matrices in a more strict way, i.e. only elements that are equal to zero are neglected. We expect that a corresponding routine that truncates the matrices with respect to a given threshold thr_{SA} would further improve the performance of the sparse LU factorization.

The error introduced by the basis functions screening and the truncation of matrices can be seen from Tab. 5.2 for the N -PDM VQMC calculations. Thus we can obtain results of 0.1 mHartree accuracy with thresholds 6/6/8 in a linear scaling time with system size for electronic structures exhibiting a significant HOMO-LUMO gap.

*CHAPTER 5. LINEAR SCALING QUANTUM MONTE CARLO ALGORITHMS FOR
THE LOCAL ENERGY WITH RIGOROUSLY CONTROLLABLE ERROR BOUNDS*

Chapter 6

Conclusions and Perspectives

In this work we have presented new linear scaling ab initio methods for predicting different properties of chemical systems. All these methods are formulated in the basis of reduced electron density matrices, which reflect the naturally local electronic structure of systems exhibiting a non-vanishing HOMO-LUMO gap. This allows the application of efficient $\mathcal{O}(M)$ sparse algebra routines since dense matrices are strictly avoided and thus an overall linear scaling behavior is provided.

Our D-GIAO-SCF method has been shown to yield an overall linear scaling behavior with system size for the calculation of NMR chemical shielding tensors at the HF and KS-DFT level of theory. While traditional $\mathcal{O}(M^3)$ methods are restricted to the hundred atoms regime, our density matrix-based method allows the treatment of systems with thousand and more atoms on single processor machines. With this at hand we are, for example, able to converge the NMR chemical shifts with respect to the influences of the surrounding environment. This has been shown for first exemplary applications to solid- and solution-state systems. The theoretical estimate of solid-state NMR shifts of a molecular host-guest complex was improved by 0.3 ppm compared to an incremental approach by calculating a pentamer containing 490 atoms with a suitable basis set (6-31G*). In a second example, we were able to converge the NMR chemical shifts of a molecular clip in solution with respect to the size of the hydration shell size, where the largest system contained 1003 atoms with a 6-31G** basis set.

In the second part of this work a new $\mathcal{O}(M)$ D-TDSCF methods for determining the linear or quadratic response to an external electric field provide an efficient way to calculate optical properties of molecular systems in a linear scaling fashion. The $\mathcal{O}(M)$ scal-

ing behavior is demonstrated for a series of linear alkanes calculated at the D-TDDFT BP86(VWN)/6-31G* level of theory. Furthermore, we describe the calculation of the first hyperpolarizabilities by exploitation of Wigner's $(2n + 1)$ rule in terms of density matrices which reduces the number of iterative calculations. The accuracy of the density matrix-based methods has been shown for the example of para-nitroaniline, where the results are compared to reference values which have been obtained in the conventional way using other established program packages. It has to be mentioned that the response to an electric field is less local than the response to a magnetic field, so one can recognize a decreased sparsity in the transition densities which results in a later onset of the linear scaling behavior^{a)}.

The new N -PDM QMC methods that cover the last part of this work provide a new approach to obtain a linear scaling behavior for the local energy in variational and diffusion quantum Monte Carlo. The exploitation of the natural sparsity of the occurring quantities with efficient sparse algebra routines has been shown to yield fast QMC results without loss of accuracy. In case of the variational approach, the reformulation in terms of a real probability density ρ_N in contrast to the wavefunction also enables one to use a fast Cholesky decomposition instead of the more involved LU algorithm. The loss of the nodal information by using the N -particle density has been prevented by reformulating the DQMC equations in terms of off-diagonal elements of the N -particle density matrix. It has been shown that the use of the sparse LU routines of the UMFPACK library [165] is not efficient for the calculated system sizes. However, because of the small prefactor of this step, we only have to resort to sparse LU factorizations for larger systems. Thus, with the possibility to account for the nodal information within our density matrix-based approach, our new method can be applied to any QMC method using trial functions like e.g. reptation QMC [169]. Furthermore it has to be mentioned that the presented schemes are of course also easily extendable to the calculation of energy derivatives [170–172].

The methods developed in this work have contributed to extend the capabilities in computing properties for large molecular systems. In this way, new insights into chemical and biochemical systems can be attained. Nevertheless, many challenges remain for methodological improvements to reduce the scaling behavior and the prefactor of algorithms to determine molecular properties or to account for electron correlation effects in large systems.

^{a)}As first calculations have shown, the number of significant elements grows with increasing frequency ω .

Bibliography

- [1] C.A. White, B.G. Johnson, P.M.W. Gill, M. Head-Gordon, *Chem. Phys. Lett.* **230**, 8 (1994).
- [2] C.A. White and M. Head-Gordon, *J. Chem. Phys.* **101**, 6593 (1994).
- [3] C.A. White, M. Head-Gordon, *J. Chem. Phys.* **105**, 5061 (1996).
- [4] C.A. White, B.G. Johnson, P.M.W. Gill, M. Head-Gordon, *Chem. Phys. Lett.*
- [5] Y. Shao, C. A. White, and M. Head-Gordon, *J. Chem. Phys.*, **114**, 6572 (2001). **253**, 268 (1996).
- [6] M. C. Strain, G. E. Scuseria, and M. J. Frisch, *Science*, **271**, 51 (1996).
- [7] M. Sierka, A. Hogekamp, and R. Ahlrichs, *J. Chem. Phys.*, **118**, 9136 (2003).
- [8] C. Ochsenfeld, C.A. White, M. Head-Gordon, *J. Chem. Phys.* **109**, 1663 (1998).
- [9] C. Ochsenfeld, *Chem. Phys. Lett.* **327**, 216 (2000).
- [10] E. Schwegler, M. Challacombe, and M. Head-Gordon, *J. Chem. Phys* **106**, 9708 (1997).
- [11] C. Ochsenfeld, M. Head-Gordon, *Chem. Phys. Lett.* **270**, 399 (1997).
- [12] M. Head-Gordon, Y. Shao, C. Saravanan, and C.A. White, *Mol. Phys.* **101**, 37 (2003).
- [13] T. Helgaker, H. Larsen, J. Olsen, P. Jørgensen, *Chem. Phys. Lett.* **327**, 397 (2000).
- [14] M. Challacombe, *J. Chem. Phys*, **110**, 2332 (1999).
- [15] J.M. Millam, G.E. Scuseria, *J. Chem. Phys.* **106**, 5569 (1997).

- [16] C. Ochsenfeld, F. Koziol, S.P. Brown, T. Schaller, U.P. Seelbach, F.-G. Klärner, *Solid State Nucl. Magn. Reson.* **22**, 128 (2002).
- [17] S.P. Brown, T. Schaller, U.P. Seelbach, F. Koziol, C. Ochsenfeld, F.-G. Klärner, H.W. Spiess, *Angew. Chemie Int. Ed.* **40**, 717 (2001).
- [18] C. Ochsenfeld, *Phys. Chem. Chem. Phys.* **2**, 2153 (2000).
- [19] C. Ochsenfeld, S.P. Brown, I. Schnell, J. Gauss, H.W. Spiess, *J. Am. Chem. Soc.* **123**, 2597 (2001).
- [20] M. Fokkens, C. Jasper T. Schrader, F. Koziol, C. Ochsenfeld, J. Polkowska, M. Lobert, B. Kahlert, F.-G. Klärner, *Chem. Eur. J.* **11**, 477 (2005).
- [21] H.-U. Siehl, T. Müller, and J. Gauss, *J. Phys. Org. Chem.* **16**, 577 (2003).
- [22] A. A. Auer, J. Gauss, and J. F. Stanton, *J. Chem. Phys.* **118**, 10407 (2003).
- [23] A. Wu, D. Cremer, and J. Gauss, *J. Phys. Chem. A* **107**, 8737 (2003).
- [24] M. Schulz-Dobrig, T. Metzroth, H. W. Spiess, J. Gauss, and I. Schnell, *Chem. Phys. Chem.* **6**, 315 (2005).
- [25] H.-U. Siehl, T. Müller, J. Gauss, P. Buzek, and P. v. R. Schleyer, *J. Am. Chem. Soc.* **116**, 6384 (1994).
- [26] Peter Günter (Ed.), *Nonlinear Optical Effects and Materials*, Springer Series in Optical Sciences Vol. 72, Springer-Verlag Berlin, 2000.
- [27] D.P. Shelton, J.E. Rice, *Chem. Rev.* **94**, 3-29 (1994).
- [28] D.M. Ceperley, L. Mitas, *Adv. Chem. Phys.* **43**, 1 (1996).
- [29] B.L. Hammond, W.A. Lester Jr., P.J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry*, World Scientific, Singapore 1994.
- [30] A. Szabo, N.S. Ostlund, *Modern Quantum Chemistry - Introduction to Advanced Electronic Structure Theory*, Dover Publications, Inc., Mineola, NY, 1989.
- [31] J.C. Slater, *Quantum Theory of Matter*, 2nd Ed., McGraw-Hill, New York, 1968.

- [32] L.D. Landau, E.M. Lifschitz, *Lehrbuch der Theoretischen Physik*, Vol. 3, *Quantenmechanik*, 9th Ed., Akademie-Verlag Berlin, 1979.
- [33] A. Messiah, *Quantum Mechanics*, Dover Publications NY, 1999.
- [34] W. Kutzelnigg, *Einführung in die Theoretische Chemie*, Bd. 1, VCH Weinheim, 1992.
- [35] C.C.J. Roothaan, *Rev. Mod. Phys.* **23**, 69 (1951).
- [36] J. Almlöf, K. Faegri, K. Korsell, *J. Comput. Chem.* **3**, 385 (1982).
- [37] M. Häser and R. Ahlrichs, *J. Comput. Chem.* **10**, 104 (1989).
- [38] S. Goedecker, G.E. Scuseria, *Comp. in Science & Engineering* **5**, 14 (2003).
- [39] L. Greengard and V. Rokhlin, *J. Comput. Phys.* **60**, 187 (1990).
- [40] Y. Shao, L. Fusti-Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. Distasio Jr., R. C. Lochan, T. Wang, G. J. O. Beran, N. A. Besley, J. M., Herbert, C. Y. Lin, T. Van Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock III, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer III, J. Kong, A. I. Krylov, P. M. W. Gill, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **8**, 3172 (2006).
- [41] R.C. Tolman, *The Principles of Statistical Mechanics*, Dover Publications NY, 1980.
- [42] R.G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, International Series of Monographies on Chemistry 16, Oxford Science Publications, 1989.
- [43] R. McWeeny, *Methods of Molecular Quantum Mechanics* (Theoretical Chemistry), 2nd Ed., Academic Press Limited, London (1989).
- [44] J. Coleman, *Rev. Mod. Phys.* **35**, 668-687 (1963).

- [45] P.A.M. Dirac, *Proc. Camb. Phil. Soc.* **27**, 240 (1931).
- [46] M. Head-Gordon, M. Lee, P. Maslen, T. van Voorhis, and S. Gwaltney, *Modern Methods and Algorithms of Quantum Chemistry*, Proceedings, 2nd Ed., J. Grotendorst (Ed.), John von Neumann Institute for Computing, Jülich, NIC Series Vol. 3, p. 593, 2000.
- [47] R. McWeeny, *Rev. Mod. Phys.* **32**, 335 (1960).
- [48] X.-P. Li, R.W. Nunes, D. Vanderbilt, *Phys. Rev. B* **47**, 10891 (1993).
- [49] P. E. Maslen, C. Ochsenfeld, C.A. White, M.S. Lee, and M. Head-Gordon, *J. Phys. Chem.* **102**, 2215 (1998).
- [50] S. Goedecker, *Rev. Mod. Phys.* **71**, 1085 (1999).
- [51] T. Kato, *Comm. Pure Appl. Math.* **10**, 151 (1957).
- [52] C. Møller, M.S. Plesset, *Phys. Rev.* **46**, 618 (1934).
- [53] J. Cížek, *J. Chem. Phys.* **45**, 4256 (1966).
- [54] J. Cížek, *Adv. Chem. Phys.* **14**, 35 (1969).
- [55] J. Cížek, J. Paldus, *Int. J. Quantum Chem.* **5**, 359 (1971).
- [56] P.R. Taylor, in *Lecture Notes in Quantum Chemistry*, Part II, B. Roos (Ed.), Springer-Verlag, Berlin, 1993.
- [57] J. Noga, R.J. Bartlett, *J. Phys. Chem.* **86**, 7041 (1987).
- [58] R.J. Bartlett, G.D. Purvis, *Int. J. Quant. Chem.* **XIV**, 561 (1978).
- [59] H. Sambe and R. H. Felton, *J. Chem. Phys.* **62**, 1122 (1975).
- [60] B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, *J. Chem. Phys.* **71**, 3396 (1979).
- [61] C. Satoko, *Chem. Phys. Lett.* **82**, 111 (1981).
- [62] R. Fournier, J. Andzelm, and D. R. Salahub, *J. Chem. Phys.* **90**, 6371 (1989).
- [63] J.A. Pople, P.W. Gill, B.G. Johnson, *Chem. Phys. Lett.* **199**, 557 (1992).

- [64] E. Teller, *Rev. Mod. Phys.* **34**, 36 (1962).
- [65] P. Hohenberg, W. Kohn, *Phys. Rev.* **136**, B864 (1964).
- [66] W. Kohn, L.J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- [67] W. Kutzelnigg, W. Klopper, *J. Chem. Phys.* **94**, 1985 (1991).
- [68] E. Runge, E.K.U. Gross, *Phys. Rev. Lett.* **52**, 997 (1984).
- [69] J.P. Perdew, A. Ruzsinszky, J. Tao, V.N. Staroverov, G.E. Scuseria, G.I. Csonka, *J. Chem. Phys.* **123**, 062201 (2005).
- [70] K. Burke, J. Werschnik, E.K.U. Gross, *J. Chem. Phys.* **123**, 062206 (2005).
- [71] F. Furche, R. Ahlrichs, *J. Chem. Phys.* **117**, 7433 (2002).
- [72] M.A.L. Marques, E.K.U. Gross, *Annu. Rev. Phys. Chem.* **55**, 427 (2004).
- [73] R. Bauernschmitt and R. Ahlrichs, *Chem. Phys. Lett.* **264**, 573 (1997).
- [74] R. van Leeuwen, *Phys. Rev. Lett.* **80**, 1280 (1998).
- [75] G.E. Scuseria, *J. Phys. Chem. A* **103**, 4782 (1999).
- [76] A.D. Becke, *J. Chem. Phys.* **98**, 1372 (1993).
- [77] P.M.W. Gill, B.G. Johnson, J.A. Pople, *Chem. Phys. Lett.* **209**, 506 (1993).
- [78] A.D. Becke, *J. Chem. Phys.* **88**, 2547 (1988).
- [79] J. Gauss, in *Modern Methods and Algorithms of Quantum Chemistry*, Proceedings, 2nd Ed., J. Grotendorst (Ed.), John von Neumann Institute for Computing, Jülich, NIC Series Vol. 3, p. 541, 2000.
- [80] C. Ochsenfeld, J. Kussmann, F. Koziol, *Angew. Chem. Int. Ed.* **43**, 4482 (2004).
- [81] P. Pulay, *Mol. Phys.* **17**, 197 (1969).
- [82] P. Pulay, in *Modern Electronic Structure Theory*, D. Yarkony (Ed.), World Scientific, Singapore, 1995, p. 1191.

- [83] P. Pulay, in *Ab Initio Methods in Quantum Chemistry*, K. P. Lawley (Ed.), John Wiley & Sons Ltd., 1987, pp. 241-286.
- [84] H. Sekino, R.J. Bartlett, *J. Chem. Phys.* **85**, 976 (1986).
- [85] J. Frenkel, *Wave Mechanics – Advanced General Theory*, Oxford University Press, Oxford 1934.
- [86] S.P. Karna, M. Dupuis, *J. Comp. Chem.* **12**, 4 (1991).
- [87] R.W. Nunes, D. Vanderbilt, *Phys. Rev. B* **50**, 17611 (1994).
- [88] G.G. Hall, *Phil. Mag.* **6**, 249 (1961).
- [89] M. Head-Gordon, P.E. Maslen and C.A. White, *J. Chem. Phys.* **108**, 616 (1998).
- [90] C.A. White, P. Maslen, M.S. Lee, M. Head-Gordon, *Chem. Phys. Lett.* **276**, 133 (1997).
- [91] H. Larsen, T. Helgaker, J. Olsen, P. Jørgensen, *J. Chem. Phys.* **115**, 10344 (2001).
- [92] M. Frisch, M. Head-Gordon, J. Pople, *Chem. Phys.* **141**, 189 (1990).
- [93] H. Larsen, J. Olsen, P. Jørgenson, and T. Helgaker, *J. Chem. Phys.* **115**, 9685 (2001).
- [94] N.F. Ramsey, *Phys. Rev.* **78**, 699 (1950).
- [95] N.F. Ramsey, *Phys. Rev.* **91**, 303 (1953).
- [96] D.P. Craig, T. Thirunamachandran, *Molecular Quantum Electrodynamics*, Dover Publications, New York, 1984.
- [97] H.F. Hamerka, *Advanced Quantum Chemistry, Theory of Interactions between Molecules and Electromagnetic Fields*, Addison-Wesley Publishing Company, Inc. Reading, 1965.
- [98] K. Wolinski, J.F. Hinton, P. Pulay, *J. Am. Chem. Soc.* **112**, 8251 (1990).
- [99] F. London, *J. Phys. Radium* **8**, 397 (1937).
- [100] R. Ditchfield, *Molecular Physics* **27**, 789 (1974).

- [101] M. Häser, R. Ahlrichs, H.P. Baron, P. Weiss, H. Horn, *Theor. Chim. Acta* **83**, 455 (1992).
- [102] T. Helgaker, M. Jaszuński, K. Ruud, *Chem. Rev.* **99**, 293 (1999).
- [103] J. Gauss, *Ber. Bunsenges. Phys. Chem.* **99**, 1001 (1995).
- [104] P. Pulay, *J. Comp. Chem.* **3**, 556 (1982).
- [105] W.H. Press et.al., *Numerical Recipes in C*, 2nd Ed., Cambridge University Press NY, 1992.
- [106] S. Obara, A. Saika, *J. Chem. Phys.* **84**, 3963 (1986);
H.B. Schlegel, *J. Chem. Phys.* **77**, 3676 (1982).
- [107] M. Head-Gordon, J.A. Pople, *J. Chem. Phys.* **89**, 5777 (1988).
- [108] F. Furche, *J. Chem. Phys.* **114**, 5982 (2001).
- [109] A.D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
- [110] J.P. Perdew, *Phys. Rev. B* **33**, 8822 (1986).
J.P. Perdew, *Phys. Rev. B* **34**, 7406 (1986) (Erratum).
J.P. Perdew, A. Zunger, *Phys. Rev. B* **23**, 5048 (1981) (Original LSDA functional).
- [111] S.H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **58**, 1200 (1980).
- [112] D.J. Thouless, *The Quantum Mechanics of Many-Body Systems*, Academic Press NY, New York 1961.
- [113] H. Weiss, R. Ahlrichs, M. Häser, *J. Chem. Phys.* **99**, 1262 (1993).
- [114] "DALTON, a molecular electronic structure program, Release 1.2 (2001)", written by T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, J. Olsen, K. Ruud, H. Ågren, A. A. Auer, K. L. Bak, V. Bakken, O. Christiansen, S. Coriani, P. Dahle, E. K. Dalskov, T. Enevoldsen, B. Fernandez, C. Hättig, K. Hald, A. Halkier, H. Heiberg, H. Hettema, D. Jonsson, S. Kirpekar, R. Kobayashi, H. Koch, K. V. Mikkelsen, P. Norman, M. J. Packer, T. B. Pedersen, T. A. Ruden, A. Sanchez, T. Saue, S. P. A. Sauer, B. Schimmelpfenning, K. O. Sylvester-Hvid, P. R. Taylor, and O. Vahtras.

- [115] P.J. Stephens, F.J. Devlin, C.F. Chabalowski, M.J. Frisch, *J. Phys. Chem.* **98**, 11623 (1994).
- [116] C. Lee, W. Yang, R.G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- [117] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, *Chem. Phys. Lett.* **162**, 165 (1989).
- [118] S.J.A. van Gisbergen, J.G. Snijders, E.J. Baerends, *J. Chem. Phys.* **109**, 10644 (1998).
- [119] S. Karna, P.N. Prasad, M. Dupuis, *J. Chem. Phys.* **94**, 1171 (1991).
- [120] P. Salek, O. Vathras, T. Helgaker, H. Ågren, *J. Chem. Phys.* **117**, 9630 (2002).
- [121] C.C. Teng, A.F. Garito, *Phys. Rev. Lett.* **50**, 350 (1983).
- [122] C.W. Murray, N.C. Handy, G.J. Laming, *Mol. Phys.* **78**, 997 (1993).
- [123] V.I. Lebedev, *Sibirsk. Mat. Zh.* **18**, 132 (1977).
V.I. Lebedev, *Zh. Vychisl. Mat. Mat. Fiz.* **15**, 48 (1975).
V.I. Lebedev, *Zh. Vychisl. Mat. Mat. Fiz.* **16**, 293 (1976).
- [124] V.I. Lebedev, D.N. Laikov, *Doklady Mathematics* **59**, 477 (1999).
- [125] F.G. Gustavson, *ACM Trans. Math. Soft* **4**, 250 (1978).
- [126] Intel MKL library, Ver. 7.2, ©1994-2004 Intel Corporation.
- [127] Sabine Schweizer, diploma thesis, AK Ochsenfeld, Universität Tübingen (2004).
- [128] Intel C/C++ compiler, Ver. 8.1, ©1996-2004 Intel Corporation.
- [129] D.M. Ceperley, G.V. Chester, M.H. Kalos, *Phys. Rev. B* **16**, 3081 (1977).
- [130] J.B. Anderson, *J. Chem. Phys.* **63**, 1499 (1975).
- [131] P.J. Reynolds, D.M. Ceperley, B.J. Alder, W.A. Lester, *J. Chem. Phys.* **77**, 5593 (1982).
- [132] S. Manten, A. Lüchow, *J. Chem. Phys.* **115**, 5362 (2001).

- [133] P. Knowles, M. Schütz, H.-J. Werner, in *Modern Methods and Algorithms of Quantum Chemistry*, Proceedings, Second Edition, J. Grotendorst (Ed.), John von Neumann Institute for Computing, Jülich, NIC Series, Vol. 3, ISBN 3-00-005834-6, p. 97-179, 2000.
- [134] P. Pulay, *Chem. Phys. Lett.* **100**, 151 (1983).
- [135] S. Manten, A. Lüchow, *J. Chem. Phys.* **119**, 1307 (2003).
- [136] D. Alfé, M.J. Gilan, *J. Phys.: Condens. Matter* **16**, L305 (2004).
- [137] A. Aspuru-Guzik, R. Slalomón-Ferrer, B. Austin, W.A. Lester Jr., *J. Comp. Chem* **26**, 708 (2005).
- [138] A. J. Williamson, R. Q. Hood and J. C. Grossmann, *Phys. Rev. Lett.* **87**, 246406 (2001).
- [139] F. A. Reboredo and A. J. Williamson, *Phys. Rev. B* **71**, 121105 (2005).
- [140] K.E. Schmidt, J.W. Moskowitz, *J. Chem. Phys.* **93**, 4172 (1990).
- [141] S.F. Boys, N.C. Handy, *Proc. R. Soc. London, Ser. A* **310**, 43 (1969).
- [142] J. Kussmann, C. Ochsenfeld, QUANTUMMC: A parallel quantum Monte Carlo code in C++ providing VQMC and DQMC algorithms.
- [143] C.W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 3rd Ed., Springer Series in Synergetics, Springer-Verlag Berlin, 2004.
- [144] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol.1, 3rd Ed., Wiley NY 1968.
- [145] H. Flyvbjerg, H.G. Petersen, *J. Chem. Phys* **91**, 416 (1989).
- [146] N. Metropolis, S. Ulam, *J. Am. Stat. Assoc.* **44**, 335 (1949).
- [147] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.M. Teller, E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).

- [148] A.A. Markov, *Extension of the limit theorems of probability theory to a sum of variables connected in a chain*, in R. Howard, *Dynamic Probabilistic Systems*, Vol. 1: *Markov Chains*, Wiley NY, 1971.
- [149] S. Fahy, X.W. Wang, S.G. Louie, *Phys. Rev. B* **42**, 3503 (1990).
- [150] M.P. Nightingale, in *Quantum Monte Carlo Methods in Physics and Chemistry*, P. Nightingale and C.J. Umrigar (Eds.), NATO ASI Series, Series C, Mathematical and Physical Sciences, Vol. 525, Kluwer Academic Publishers, Boston, 1999.
- [151] C.J. Umrigar, in *Quantum Monte Carlo Methods in Physics and Chemistry*, P. Nightingale and C.J. Umrigar (Eds.), NATO ASI Series, Series C, Mathematical and Physical Sciences, Vol. 525, Kluwer Academic Publishers, Boston, 1999.
- [152] C.J. Umrigar, K.G. Wilson, J.W. Wilkins, *Phys. Rev. Lett.* **60**, 1719 (1988).
- [153] C.J. Umrigar, M.P. Nightingale, K.J. Runge, *J. Chem. Phys.* **99**, 2865 (1993).
- [154] H.F. Trotter, *Proc. Am. Math. Soc.* **10**, 545 (1959).
- [155] M. Suzuki, S. Miyashita, A. Kuroda, *Prog. Theor. Phys.* **58**, 1377 (1977).
- [156] R.C. Grimm, R.G. Storer, *J. Comp. Phys.* **7**, 134 (1971).
- [157] W.M.C. Foulkes, L. Mitas, R.J. Needs, G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- [158] L. Mitas, in *Quantum Monte Carlo Methods in Physics and Chemistry*, P. Nightingale and C.J. Umrigar (Eds.), NATO ASI Series, Series C, Mathematical and Physical Sciences, Vol. 525, Kluwer Academic Publishers, Boston, 1999.
- [159] Hella Riede, diploma thesis, AK Ochsenfeld, Universität Tübingen (2005).
- [160] C.R. Myers, C.J. Umrigar, J.P. Sethna, J.D. Morgan III, *Phys. Rev. A* **44**, 5537 (1991).
- [161] S.F. Boys, *Rev. Mod. Phys.* **32**, 296 (1960).
- [162] C. Edmiston, K. Ruedenberg, *J. Chem. Phys.* **43**, 97 (1965).
- [163] J. Pipek, P.H. Mezey, *J. Chem. Phys.* **90**, 4916 (1989).

-
- [164] T.A. Davies, I.S. Duff, *SIAM J. Matrix Anal. Appl.* **18**, 140 (1997).
- [165] UMFPACK Version 4.4 (Jan. 28, 2005), Copyright © 2005 by Timothy A. Davis.
- [166] T.H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
- [167] D.M. Ceperley, *Rev. Mod. Phys.* **67**, 279 (1995).
- [168] D.M. Ceperley, in *Simulation in Condensed Matter Physics and Chemistry*, K. Binder and G. Ciccotti (Ed.), 1996.
- [169] S. Baroni, S. Moroni, *Phys. Rev. Lett.* **82**, 4745 (1999).
- [170] R. Assaraf, M. Caffarel, *J. Chem. Phys.* **119**, 10536 (2003).
- [171] C.J. Umrigar, *Int. J. Quantum Chem.* **23**, 217 (1989).
- [172] C. Filippi, C.J. Umrigar, *Phys. Rev. B* **61**, R16291 (2000).

List of Figures

2.1	McWeeny's purification function.	22
2.2	Sparsity pattern of the MO coefficient matrix and one-electron density. . .	23
2.3	Exchange and Coulomb hole.	26
2.4	Hierarchy of correlation methods.	27
3.1	Pentamer of host-guest complex.	68
3.2	<i>N</i> -methyl nicotinamide in hydration shell.	69
3.3	Matrix Patterns of $\epsilon^x(\pm\omega)$ and $\mathbf{U}^x(\pm\omega)$	81
3.4	Multiplication of Matrix Patterns	82
3.5	p-Nitroaniline	84
4.1	Performance of sparse matrix multiplications.	95
5.1	<i>N</i> -PDM VQMC calculations of a series of methane clusters.	117
5.2	<i>N</i> -PDM VQMC calculations of a series of linear alkanes: total CPU times. . .	119
5.3	<i>N</i> -PDM FN-DQMC calculations of a series of linear alkanes: total CPU times.	125
5.4	<i>N</i> -PDM FN-DQMC calculations of a series of linear alkanes: CPU times for the determination of the inverse $\tilde{\mathbf{D}}_\rho^\dagger$ and the determinant value by LU factorization.	126
B.1	Convergence of McWeeny's purification transformation.	155
B.2	Localization of electrons at the example of Peierls distortion.	156
B.3	Timings for GIAO-HF/6-31G* of a series of linear alkanes.	157
B.4	Timings for GIAO-KS-DFT BP86(VWN5)/6-31G* of a series of 1-4 con- nected amylose chains.	158
B.5	Scheme of D-CPSCF algorithm.	159

B.6	Timings for TDDFT-BP86(VWN)/6-31G* to calculate the static polarizability tensor of a series of linear alkanes.	160
B.7	Frequency-dependent polarizability for hydrogen fluoride with BP86(VWN)/6-31G*.	161
B.8	Timings for TDDFT-BP86(VWN)/6-31G* using Wigner's $(2n + 1)$ rule to calculate the static hyperpolarizability tensor of a series of linear alkanes.	162
B.9	Fluctuations in the maximal element of \mathbf{P} within 1000 purification steps of a HCF ₃ molecule with HF/6-31G*.	163
B.10	Flop-counting for 1000 purification steps of a HCF ₃ molecule with HF/6-31G* for different screening schemes.	164
B.11	Sparsity in percent of \mathbf{P} within 1000 purification steps of a HCF ₃ molecule with HF/6-31G* for different screening schemes.	165
B.12	Sum of Giga-Flops for different Level-1 iterations within GIAO-KS-DFT BP86(VWN)/6-31G* of an amylose chain containing 8 a-D-glucose units.	166
B.13	Comparison of screened and unscreened sparse matrix multiplication at the example of an amylose chain containing 8 a-D-glucose units.	167
B.14	Scheme of sparse matrix multiplications.	168
B.15	Cusp-correction of 1s-function of a lithium atom (cc-pVTZ basis set).	169
B.16	N -PDM VQMC: scheme to compute ρ_N and its 1st and 2nd derivatives with respect to the single electron positions.	170
B.17	N -PDM VQMC calculations of a series of linear alkanes: CPU times for basis functions evaluation.	171
B.18	N -PDM VQMC calculations of a series of linear alkanes: CPU times for matrix multiplies.	172
B.19	N -PDM VQMC calculations of a series of linear alkanes: number of significant elements in \mathbf{P} for different thresholds.	173
B.20	N -PDM FN-DQMC: scheme to compute ρ_N and its 1st and 2nd derivatives with respect to the single electron positions.	174
B.21	N -PDM FN-DQMC calculations of a series of linear alkanes: total CPU times using the sparse LU factorization of the UMFPACK.	175
B.22	N -PDM FN-DQMC calculations of a series of linear alkanes: total CPU times using the sparse LU factorization of the UMFPACK after recompression of the determinant matrix.	176

List of Tables

3.1	Polarizabilities and first order hyperpolarizabilities.	40
3.2	Maximal change of NMR shifts of the host-guest system with increasing hydration shell.	70
3.3	Excitation energies and oscillator strengths for the first three excited singlet states from a TDDFT calculation with BP86(VWN)/6-31G* of HCF ₃ . . .	74
3.4	Relative oscillator strengths of the three lowest excitations of HCF ₃ with BP86(VWN)/6-31G*.	75
3.5	Dynamic polarizabilities of PNA with TDHF/6-31G* and TDDFT-B3LYP/6-31G*.	76
3.6	Second harmonic generation for PNA with B3LYP/6-31G*.	85
4.1	CPU timings for 38 Level-1 iterations of an amylose chain containing 8 α -D-glucose units with different sparse multiplication screening schemes. . . .	90
4.2	Stability analysis of sparse matrix multiplications at the example of HCF ₃ with HF/6-31G*.	92
4.3	Effect of truncating matrices for several systems with HF/6-31G* at the example of calculating $\mathbf{S}^{-1}\mathbf{F}$	93
4.4	Effect of truncating matrices for several systems with HF/6-31G* at the example of calculating \mathbf{PS}	93
4.5	Comparison of sparse matrix multiplications with a standard $\mathcal{O}(M^3)$ routine.	94
5.1	N -PDM VQMC calculations of a series of linear alkanes: scaling behavior of the number of significant elements in the one-electron density matrix \mathbf{P}	120
5.2	N -PDM VQMC calculations of a series of linear alkanes: error estimation via correlated sampling.	121

C.1	Structure of para-nitroaniline.	177
C.2	Frequency-dependent polarizabilities of para-nitroaniline with HF/6-31G*.	178
C.3	Frequency-dependent polarizabilities of para-nitroaniline with B3LYP/6-31G*.	178
C.4	Frequency-dependent first hyperpolarizabilities of para-nitroaniline with HF/6-31G*.	179
C.5	Frequency-dependent first hyperpolarizabilities of para-nitroaniline with B3LYP/6-31G*.	179
C.6	Second harmonic generation of para-nitroaniline with D-TDDFT B3LYP/6-31G*.	180
C.7	Explicit sparsity pattern of \mathbf{P} after 1000 purification steps of HCF ₃ molecule with HF/6-31G*.	180
C.8	Convergence behavior of GIAO-HF/6-31G* at the examples of a series of linear alkanes.	181
C.9	Integer coefficients of EJ ₅ short-ranged Schmidt-Moskowitz correlation factor [135].	181

Appendix A

Derivatives of the Exchange-Correlation Potential

In the following the general equations to form the third order derivatives of the XC energy functional of LSDA and GGA type, respectively, are derived. The partial derivatives of the functionals to form the final term in eq. 3.175 as well as the first order transition densities have to be provided. In this work the third order derivatives of the following exchange and correlation functionals have been implemented:

Exchange	Correlation
Slater's $X\alpha$	Wigner
Becke (1988)	Vosko/Wilks/Nusair (RPA)
	Vosko/Wilks/Nusair (No. 5)
	Perdew-Zunger (1981)
	Perdew's GGA (1986, with VWN (No. 5) and Perdew-Zunger LSDA kernel)
	Lee-Yang-Parr

So the popular Becke's three-parameter formula in combination with the LYP functional is also possible.

A.1 Local Spin Density Approximation

Since the LSDA functionals only depend on the local density $E_{xc} = f(\rho)$, the XC potential and its first and second order derivatives are easily derived as

$$\langle \chi_\mu | \hat{v}_{xc} | \chi_\nu \rangle_\alpha = \int \frac{\partial E_{xc}}{\partial \rho_\alpha} \chi_\mu \chi_\nu d\mathbf{r}, \quad (\text{A.1})$$

$$\langle \chi_\mu | \hat{v}_{xc}^x | \chi_\nu \rangle_\alpha = \int \left(\frac{\partial^2 E_{xc}}{\partial \rho_\alpha^2} \rho_\alpha^x + \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta} \rho_\beta^x \right) \chi_\mu \chi_\nu d\mathbf{r}, \quad (\text{A.2})$$

$$\begin{aligned} \langle \chi_\mu | \hat{v}_{xc}^{xy} | \chi_\nu \rangle_\alpha &= \int \left(\frac{\partial^2 E_{xc}}{\partial \rho_\alpha^2} \rho_\alpha^{xy} + \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta} \rho_\beta^{xy} \right) \chi_\mu \chi_\nu d\mathbf{r} \\ &+ \int \left(\frac{\partial^3 E_{xc}}{\partial \rho_\alpha^3} \rho_\alpha^x \rho_\alpha^y + \frac{\partial^3 E_{xc}}{\partial \rho_\alpha^2 \partial \rho_\beta} [\rho_\alpha^x \rho_\beta^y + \rho_\alpha^y \rho_\beta^x] \right. \\ &\left. + \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta^2} \rho_\beta^x \rho_\beta^y \right) \chi_\mu \chi_\nu d\mathbf{r}, \end{aligned} \quad (\text{A.3})$$

with similar expressions for β . As mentioned before, the last term of the second derivative can be directly computed from the first order transition densities ρ^x while the first term is treated in the iterative process with the same routines as used for the linear response in eq. A.2.

A.2 Generalized Gradient Approximation

The case for GGA functionals is more complicated since the XC functional also depends on the density gradient $|\nabla\rho|$. The XC potential in the given basis is

$$\langle \chi_\mu | \hat{v}_{xc} | \chi_\nu \rangle = \int \frac{\partial E_{xc}}{\partial \rho} \chi_\mu \chi_\nu d\mathbf{r} + \int \frac{\partial E_{xc}}{\partial |\nabla\rho|} \frac{\nabla_i \rho}{|\nabla\rho|} \nabla_i (\chi_\mu \chi_\nu), \quad (\text{A.4})$$

where the Einstein sum convention is used for the index i of the gradient components of the density and basis functions. Within a computational scheme it is easier to treat the functional derivatives with respect to the square of the density gradient $\zeta = (\nabla\rho)^2$, so with the substitution

$$\begin{aligned} \frac{\partial E_{xc}}{\partial |\nabla\rho|} &= \frac{\partial E_{xc}}{\partial \zeta} \frac{\partial \zeta}{\partial |\nabla\rho|} \\ &= \frac{\partial E_{xc}}{\partial \zeta} 2|\nabla\rho| \end{aligned} \quad (\text{A.5})$$

APPENDIX A. DERIVATIVES OF THE EXCHANGE-CORRELATION POTENTIAL

we obtain for eq. A.4

$$\langle \chi_\mu | \hat{v}_{xc} | \chi_\nu \rangle = \int \frac{\partial E_{xc}}{\partial \rho} \chi_\mu \chi_\nu d\mathbf{r} + 2 \int \frac{\partial E_{xc}}{\partial \zeta} \nabla_i \rho \nabla_i (\chi_\mu \chi_\nu). \quad (\text{A.6})$$

From the functional evaluation we obtain the derivatives with respect to ρ and $(\nabla_i \rho)^2$, so we form the linear response of the XC potential as

$$\begin{aligned} \langle \chi_\mu | \hat{v}_{xc}^x | \chi_\nu \rangle_\alpha &= \int \left(\frac{\partial^2 E_{xc}}{\partial \rho_\alpha^2} \rho_\alpha^x + \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta} \rho_\beta^x \right. \\ &+ 2 \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x + 2 \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \zeta_\beta} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \\ &+ \left. \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \zeta_{\alpha\beta}} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \right) \chi_\mu \chi_\nu d\mathbf{r} \\ &+ \int \left(4 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha^2} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x + 4 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha \partial \zeta_\beta} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \right. \\ &+ 2 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha \partial \zeta_{\alpha\beta}} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \\ &+ \left. 2 \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha} \rho_\alpha^x + 2 \frac{\partial^2 E_{xc}}{\partial \rho_\beta \partial \zeta_\alpha} \rho_\beta^x \right) \nabla_j \rho_\alpha \nabla_j (\chi_\mu \chi_\nu) d\mathbf{r} \\ &+ \int \left(2 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha \partial \zeta_{\alpha\beta}} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x + 2 \frac{\partial^2 E_{xc}}{\partial \zeta_\beta \partial \zeta_{\alpha\beta}} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \right. \\ &+ 2 \frac{\partial^2 E_{xc}}{\partial \zeta_{\alpha\beta}^2} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \\ &+ \left. \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \zeta_{\alpha\beta}} \rho_\alpha^x + \frac{\partial^2 E_{xc}}{\partial \rho_\beta \partial \zeta_{\alpha\beta}} \rho_\beta^x \right) \nabla_j \rho_\beta \nabla_j (\chi_\mu \chi_\nu) d\mathbf{r} \\ &+ \int \left(2 \frac{\partial E_{xc}}{\partial \zeta_\alpha} \nabla_i \rho_\alpha^x + \frac{\partial E_{xc}}{\partial \zeta_{\alpha\beta}} \nabla_i \rho_\beta^x \right) \nabla_i (\chi_\mu \chi_\nu) d\mathbf{r}. \quad (\text{A.7}) \end{aligned}$$

While denoting $\langle \chi_\mu | \hat{v}_{xc}^x(\rho^{xy}) | \chi_\nu \rangle_\alpha$ as eq. A.7 with ρ^x replaced by ρ^{xy} , a further differentiation yields the quadratic response. Since the expression is a bit lengthy, we split it up into a sum of terms presented in the following:

$$\begin{aligned} \langle \chi_\mu | \hat{v}_{xc}^{xy} | \chi_\nu \rangle_\alpha &= \langle \chi_\mu | \hat{v}_{xc}^x(\rho^{xy}) | \chi_\nu \rangle_\alpha \\ &+ T_1 + T_2 + T_3 + T_4 + T_5. \quad (\text{A.8}) \end{aligned}$$

$$\begin{aligned}
T_1 = & \int \left(2 \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha} \nabla_i \rho_\alpha^y \nabla_i \rho_\alpha^x + 2 \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \zeta_\beta} \nabla_i \rho_\beta^y \nabla_i \rho_\beta^x \right. \\
& + \left. \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \zeta_{\alpha\beta}} [\nabla_i \rho_\alpha^y \nabla_i \rho_\beta^x + \nabla_i \rho_\beta^y \nabla_i \rho_\alpha^x] \right) \chi_\mu \chi_\nu d\mathbf{r} \\
& + \int \left(4 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha^2} \nabla_i \rho_\alpha^y \nabla_i \rho_\alpha^x + 4 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha \partial \zeta_\beta} \nabla_i \rho_\beta^y \nabla_i \rho_\beta^x \right. \\
& + 2 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha \partial \zeta_{\alpha\beta}} [\nabla_i \rho_\alpha^y \nabla_i \rho_\beta^x + \nabla_i \rho_\beta^y \nabla_i \rho_\alpha^x] \\
& + \left. 2 \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha} \rho_\alpha^x + 2 \frac{\partial^2 E_{xc}}{\partial \rho_\beta \partial \zeta_\alpha} \rho_\beta^x \right) \nabla_j \rho_\alpha \nabla_j (\chi_\mu \chi_\nu) d\mathbf{r} \\
& + \int \left(4 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha^2} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x + 4 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha \partial \zeta_\beta} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \right. \\
& + 2 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha \partial \zeta_{\alpha\beta}} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \\
& + \left. 2 \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha} \rho_\alpha^x + 2 \frac{\partial^2 E_{xc}}{\partial \rho_\beta \partial \zeta_\alpha} \rho_\beta^x \right) \nabla_j \rho_\alpha^y \nabla_j (\chi_\mu \chi_\nu) d\mathbf{r} \\
& + \int \left(2 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha \partial \zeta_{\alpha\beta}} \nabla_i \rho_\alpha^y \nabla_i \rho_\alpha^x + 2 \frac{\partial^2 E_{xc}}{\partial \zeta_\beta \partial \zeta_{\alpha\beta}} \nabla_i \rho_\beta^y \nabla_i \rho_\beta^x \right. \\
& + 2 \frac{\partial^2 E_{xc}}{\partial \zeta_{\alpha\beta}^2} [\nabla_i \rho_\alpha^y \nabla_i \rho_\beta^x + \nabla_i \rho_\beta^y \nabla_i \rho_\alpha^x] \\
& + \left. \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \zeta_{\alpha\beta}} \rho_\alpha^x + \frac{\partial^2 E_{xc}}{\partial \rho_\beta \partial \zeta_{\alpha\beta}} \rho_\beta^x \right) \nabla_j \rho_\beta \nabla_j (\chi_\mu \chi_\nu) d\mathbf{r} \\
& + \int \left(2 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha \partial \zeta_{\alpha\beta}} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x + 2 \frac{\partial^2 E_{xc}}{\partial \zeta_\beta \partial \zeta_{\alpha\beta}} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \right. \\
& + 2 \frac{\partial^2 E_{xc}}{\partial \zeta_{\alpha\beta}^2} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \\
& + \left. \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \zeta_{\alpha\beta}} \rho_\alpha^x + \frac{\partial^2 E_{xc}}{\partial \rho_\beta \partial \zeta_{\alpha\beta}} \rho_\beta^x \right) \nabla_j \rho_\beta^y \nabla_j (\chi_\mu \chi_\nu) d\mathbf{r}
\end{aligned} \tag{A.9}$$

$$\begin{aligned}
T_2 = & \int \left(\frac{\partial^3 E_{xc}}{\partial \rho_\alpha^3} \rho_\alpha^x \rho_\alpha^y + \frac{\partial^3 E_{xc}}{\partial \rho_\alpha^2 \partial \rho_\beta} [\rho_\alpha^x \rho_\beta^y + \rho_\alpha^y \rho_\beta^x] + \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta^2} \rho_\beta^x \rho_\beta^y \right. \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha^2 \partial \zeta_\alpha} \rho_\alpha^x \nabla_i \rho_\alpha \nabla_i \rho_\alpha^y + \frac{\partial^3 E_{xc}}{\partial \rho_\alpha^2 \partial \zeta_{\alpha\beta}} \rho_\alpha^x [\nabla_i \rho_\alpha \nabla_i \rho_\beta^y + \nabla_i \rho_\beta \nabla_i \rho_\alpha^y] \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha^2 \partial \zeta_\beta} \rho_\alpha^x \nabla_i \rho_\beta \nabla_i \rho_\beta^y + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta \partial \zeta_\alpha} \rho_\beta^x \nabla_i \rho_\alpha \nabla_i \rho_\alpha^y \\
& + \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta \partial \zeta_{\alpha\beta}} \rho_\beta^x [\nabla_i \rho_\alpha \nabla_i \rho_\beta^y + \nabla_i \rho_\beta \nabla_i \rho_\alpha^y] \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta \partial \zeta_\beta} \rho_\beta^x \nabla_i \rho_\beta \nabla_i \rho_\beta^y + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha^2 \partial \zeta_\alpha} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x \rho_\alpha^y \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha^2 \partial \zeta_\beta} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \rho_\alpha^y + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha^2 \partial \zeta_{\alpha\beta}} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \rho_\alpha^y \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta \partial \zeta_\alpha} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x \rho_\beta^y + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta \partial \zeta_\beta} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \rho_\beta^y \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta \partial \zeta_{\alpha\beta}} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \rho_\beta^y + 4 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha^2} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x \nabla_j \rho_\alpha \nabla_j \rho_\alpha^y \\
& + 4 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha \partial \zeta_\beta} [\nabla_i \rho_\beta \nabla_i \rho_\beta^x \nabla_j \rho_\alpha \nabla_j \rho_\alpha^y + \nabla_i \rho_\beta \nabla_i \rho_\beta^y \nabla_j \rho_\alpha \nabla_j \rho_\alpha^x] \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha \partial \zeta_{\alpha\beta}} [[\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \nabla_j \rho_\alpha \nabla_j \rho_\alpha^y \\
& + [\nabla_i \rho_\alpha \nabla_i \rho_\beta^y + \nabla_i \rho_\beta \nabla_i \rho_\alpha^y] \nabla_j \rho_\alpha \nabla_j \rho_\alpha^x] \\
& + \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_{\alpha\beta}^2} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] [\nabla_j \rho_\alpha \nabla_j \rho_\beta^y + \nabla_j \rho_\beta \nabla_j \rho_\alpha^y] \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\beta \partial \zeta_{\alpha\beta}} [[\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \nabla_j \rho_\beta \nabla_j \rho_\beta^y \\
& + [\nabla_i \rho_\alpha \nabla_i \rho_\beta^y + \nabla_i \rho_\beta \nabla_i \rho_\alpha^y] \nabla_j \rho_\beta \nabla_j \rho_\beta^x] \\
& \left. + 4 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\beta^2} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \nabla_j \rho_\beta \nabla_j \rho_\beta^y \right) \chi_\mu \chi_\nu d\mathbf{r} \tag{A.10}
\end{aligned}$$

$$\begin{aligned}
 T_3 = & \int \left(2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha^2 \partial \zeta_\alpha} \rho_\alpha^x \rho_\alpha^y + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta \partial \zeta_\alpha} \rho_\alpha^x \rho_\beta^y \right. \\
 & + 4 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha^2} \rho_\alpha^x \nabla_i \rho_\alpha \nabla_i \rho_\alpha^y + 4 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha \partial \zeta_\beta} \rho_\alpha^x \nabla_i \rho_\alpha \nabla_i \rho_\beta^y \\
 & + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha \partial \zeta_{\alpha\beta}} \rho_\alpha^x [\nabla_i \rho_\alpha \nabla_i \rho_\beta^y + \nabla_i \rho_\beta \nabla_i \rho_\alpha^y] \\
 & + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta \partial \zeta_\alpha} \rho_\beta^x \rho_\alpha^y + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\beta^2 \partial \zeta_\alpha} \rho_\beta^x \rho_\beta^y \\
 & + 4 \frac{\partial^3 E_{xc}}{\partial \rho_\beta \partial \zeta_\alpha^2} \rho_\beta^x \nabla_i \rho_\alpha \nabla_i \rho_\alpha^y + 4 \frac{\partial^3 E_{xc}}{\partial \rho_\beta \partial \zeta_\alpha \partial \zeta_\beta} \rho_\beta^x \nabla_i \rho_\beta \nabla_i \rho_\alpha^y \\
 & + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\beta \partial \zeta_\alpha \partial \zeta_{\alpha\beta}} \rho_\beta^x [\nabla_i \rho_\alpha \nabla_i \rho_\beta^y + \nabla_i \rho_\beta \nabla_i \rho_\alpha^y] \\
 & + 4 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha^2} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x \rho_\alpha^y + 4 \frac{\partial^3 E_{xc}}{\partial \rho_\beta \partial \zeta_\alpha^2} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x \rho_\beta^y \\
 & + 8 \frac{\partial^3 E_{xc}}{\partial \zeta_\alpha^3} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x \nabla_j \rho_\alpha \nabla_j \rho_\alpha^y + 8 \frac{\partial^3 E_{xc}}{\partial \zeta_\alpha^2 \partial \zeta_\beta} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x \nabla_j \rho_\beta \nabla_j \rho_\beta^y \\
 & + 4 \frac{\partial^3 E_{xc}}{\partial \zeta_\alpha^2 \partial \zeta_{\alpha\beta}} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x [\nabla_j \rho_\alpha \nabla_j \rho_\beta^y + \nabla_j \rho_\beta \nabla_j \rho_\alpha^y] \\
 & + 4 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha \partial \zeta_\beta} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \rho_\alpha^y + 4 \frac{\partial^3 E_{xc}}{\partial \rho_\beta \partial \zeta_\alpha \partial \zeta_\beta} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \rho_\beta^y \\
 & + 8 \frac{\partial^3 E_{xc}}{\partial \zeta_\alpha^2 \partial \zeta_\beta} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \nabla_j \rho_\alpha \nabla_j \rho_\alpha^y + 8 \frac{\partial^3 E_{xc}}{\partial \zeta_\alpha \partial \zeta_\beta^2} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \nabla_j \rho_\beta \nabla_j \rho_\beta^y \\
 & + 4 \frac{\partial^3 E_{xc}}{\partial \zeta_\alpha \partial \zeta_\beta \partial \zeta_{\alpha\beta}} \nabla_i \rho_\beta \nabla_i \rho_\beta^x [\nabla_j \rho_\alpha \nabla_j \rho_\beta^y + \nabla_j \rho_\beta \nabla_j \rho_\alpha^y] \\
 & + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha \partial \zeta_{\alpha\beta}} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \rho_\alpha^y \\
 & + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\beta \partial \zeta_\alpha \partial \zeta_{\alpha\beta}} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \rho_\beta^y \\
 & + 4 \frac{\partial^3 E_{xc}}{\partial \zeta_\alpha^2 \partial \zeta_{\alpha\beta}} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \nabla_j \rho_\alpha \nabla_j \rho_\alpha^y \\
 & + 4 \frac{\partial^3 E_{xc}}{\partial \zeta_\alpha \partial \zeta_\beta \partial \zeta_{\alpha\beta}} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \nabla_j \rho_\beta \nabla_j \rho_\beta^y \\
 & + 2 \frac{\partial^3 E_{xc}}{\partial \zeta_\alpha \partial \zeta_{\alpha\beta}^2} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] [\nabla_j \rho_\alpha \nabla_j \rho_\beta^y + \nabla_j \rho_\beta \nabla_j \rho_\alpha^y] \\
 & \left. \right) \nabla_k \rho_\alpha \nabla_k (\chi_\mu \chi_\nu) d\mathbf{r}. \tag{A.11}
 \end{aligned}$$

$$\begin{aligned}
T_4 = & \int \left(\frac{\partial^3 E_{xc}}{\partial \rho_\alpha^2 \partial \zeta_{\alpha\beta}} \rho_\alpha^x \rho_\alpha^y + \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta \partial \zeta_{\alpha\beta}} \rho_\alpha^x \rho_\beta^y \right. \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha \partial \zeta_{\alpha\beta}} \rho_\alpha^x \nabla_i \rho_\alpha \nabla_i \rho_\alpha^y + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\beta \partial \zeta_{\alpha\beta}} \rho_\alpha^x \nabla_i \rho_\beta \nabla_i \rho_\beta^y \\
& + \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_{\alpha\beta}^2} \rho_\alpha^x [\nabla_i \rho_\alpha \nabla_i \rho_\beta^y + \nabla_i \rho_\beta \nabla_i \rho_\alpha^y] \\
& + \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \rho_\beta \partial \zeta_{\alpha\beta}} \rho_\beta^x \rho_\alpha^y + \frac{\partial^3 E_{xc}}{\partial \rho_\beta^2 \partial \zeta_{\alpha\beta}} \rho_\beta^x \rho_\beta^y \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\beta \partial \zeta_\alpha \partial \zeta_{\alpha\beta}} \rho_\beta^x \nabla_i \rho_\alpha \nabla_i \rho_\alpha^y + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\beta \partial \zeta_\beta \partial \zeta_{\alpha\beta}} \rho_\beta^x \nabla_i \rho_\beta \nabla_i \rho_\beta^y \\
& + \frac{\partial^3 E_{xc}}{\partial \rho_\beta \partial \zeta_{\alpha\beta}^2} \rho_\beta^x [\nabla_i \rho_\alpha \nabla_i \rho_\beta^y + \nabla_i \rho_\beta \nabla_i \rho_\alpha^y] \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha \partial \zeta_{\alpha\beta}} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x \rho_\alpha^y + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\beta \partial \zeta_\alpha \partial \zeta_{\alpha\beta}} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x \rho_\beta^y \\
& + 4 \frac{\partial^3 E_{xc}}{\partial \zeta_\alpha^2 \partial \zeta_{\alpha\beta}} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x \nabla_j \rho_\alpha \nabla_j \rho_\alpha^y + 4 \frac{\partial^3 E_{xc}}{\partial \zeta_\alpha \partial \zeta_\beta \partial \zeta_{\alpha\beta}} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x \nabla_j \rho_\beta \nabla_j \rho_\beta^y \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \zeta_\alpha \partial \zeta_{\alpha\beta}^2} \nabla_i \rho_\alpha \nabla_i \rho_\alpha^x [\nabla_j \rho_\alpha \nabla_j \rho_\beta^y + \nabla_j \rho_\beta \nabla_j \rho_\alpha^y] \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_\beta \partial \zeta_{\alpha\beta}} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \rho_\alpha^y + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\beta \partial \zeta_\beta \partial \zeta_{\alpha\beta}} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \rho_\beta^y \\
& + 4 \frac{\partial^3 E_{xc}}{\partial \zeta_\alpha \partial \zeta_\beta \partial \zeta_{\alpha\beta}} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \nabla_j \rho_\alpha \nabla_j \rho_\alpha^y + 4 \frac{\partial^3 E_{xc}}{\partial \zeta_\beta^2 \partial \zeta_{\alpha\beta}} \nabla_i \rho_\beta \nabla_i \rho_\beta^x \nabla_j \rho_\beta \nabla_j \rho_\beta^y \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \zeta_\beta \partial \zeta_{\alpha\beta}^2} \nabla_i \rho_\beta \nabla_i \rho_\beta^x [\nabla_j \rho_\alpha \nabla_j \rho_\beta^y + \nabla_j \rho_\beta \nabla_j \rho_\alpha^y] \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\alpha \partial \zeta_{\alpha\beta}^2} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \rho_\alpha^y \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \rho_\beta \partial \zeta_{\alpha\beta}^2} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \rho_\beta^y \\
& + 4 \frac{\partial^3 E_{xc}}{\partial \zeta_\alpha \partial \zeta_{\alpha\beta}^2} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \nabla_j \rho_\alpha \nabla_j \rho_\alpha^y \\
& + 4 \frac{\partial^3 E_{xc}}{\partial \zeta_\beta \partial \zeta_{\alpha\beta}^2} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] \nabla_j \rho_\beta \nabla_j \rho_\beta^y \\
& + 2 \frac{\partial^3 E_{xc}}{\partial \zeta_{\alpha\beta}^3} [\nabla_i \rho_\alpha \nabla_i \rho_\beta^x + \nabla_i \rho_\beta \nabla_i \rho_\alpha^x] [\nabla_j \rho_\alpha \nabla_j \rho_\beta^y + \nabla_j \rho_\beta \nabla_j \rho_\alpha^y] \\
& \left. \right) \nabla_k \rho_\beta \nabla_k (\chi_\mu \chi_\nu) \tag{A.12}
\end{aligned}$$

APPENDIX A. DERIVATIVES OF THE EXCHANGE-CORRELATION POTENTIAL

$$\begin{aligned}
T_5 = & \int \left(2 \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \zeta_\alpha} \nabla_i \rho_\alpha^x \rho_\alpha^y + 2 \frac{\partial^2 E_{xc}}{\partial \rho_\beta \partial \zeta_\alpha} \nabla_i \rho_\alpha^x \rho_\beta^y \right. \\
& + 4 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha^2} \nabla_i \rho_\alpha^x \nabla_j \rho_\alpha^y + 4 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha \partial \zeta_\beta} \nabla_i \rho_\alpha^x \nabla_j \rho_\beta^y \nabla_j \rho_\beta^y \\
& + 2 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha \partial \zeta_{\alpha\beta}} \nabla_i \rho_\alpha^x [\nabla_j \rho_\alpha^y \nabla_j \rho_\beta^y + \nabla_j \rho_\beta^y \nabla_j \rho_\alpha^y] \\
& + \frac{\partial^2 E_{xc}}{\partial \rho_\alpha \partial \zeta_{\alpha\beta}} \nabla_i \rho_\beta^x \rho_\alpha^y + \frac{\partial^2 E_{xc}}{\partial \rho_\beta \partial \zeta_{\alpha\beta}} \nabla_i \rho_\beta^x \rho_\beta^y \\
& + 2 \frac{\partial^2 E_{xc}}{\partial \zeta_\alpha \partial \zeta_{\alpha\beta}} \nabla_i \rho_\beta^x \nabla_j \rho_\alpha^y \nabla_j \rho_\alpha^y + 2 \frac{\partial^2 E_{xc}}{\partial \zeta_\beta \partial \zeta_{\alpha\beta}} \nabla_i \rho_\beta^x \nabla_j \rho_\beta^y \nabla_j \rho_\beta^y \\
& \left. + \frac{\partial^2 E_{xc}}{\partial \zeta_{\alpha\beta}^2} \nabla_i \rho_\beta^x [\nabla_j \rho_\alpha^y \nabla_j \rho_\beta^y + \nabla_j \rho_\beta^y \nabla_j \rho_\alpha^y] \right) \nabla_i (\chi_\mu \chi_\nu)
\end{aligned} \tag{A.13}$$

Appendix B

Figures

Figure B.1: Convergence of purification transformation for different starting values (left). Purification of density of α -D-glucose within numerical accuracy ($\sim 10^{-16}$) after one geometry optimization step within D-QCSCF calculation with HF/6-31G* (right), the logarithmic value of the norm of the residual ($\log \|\mathbf{P}_i - \mathbf{P}_{i-1}\|$) is given as test for convergence.

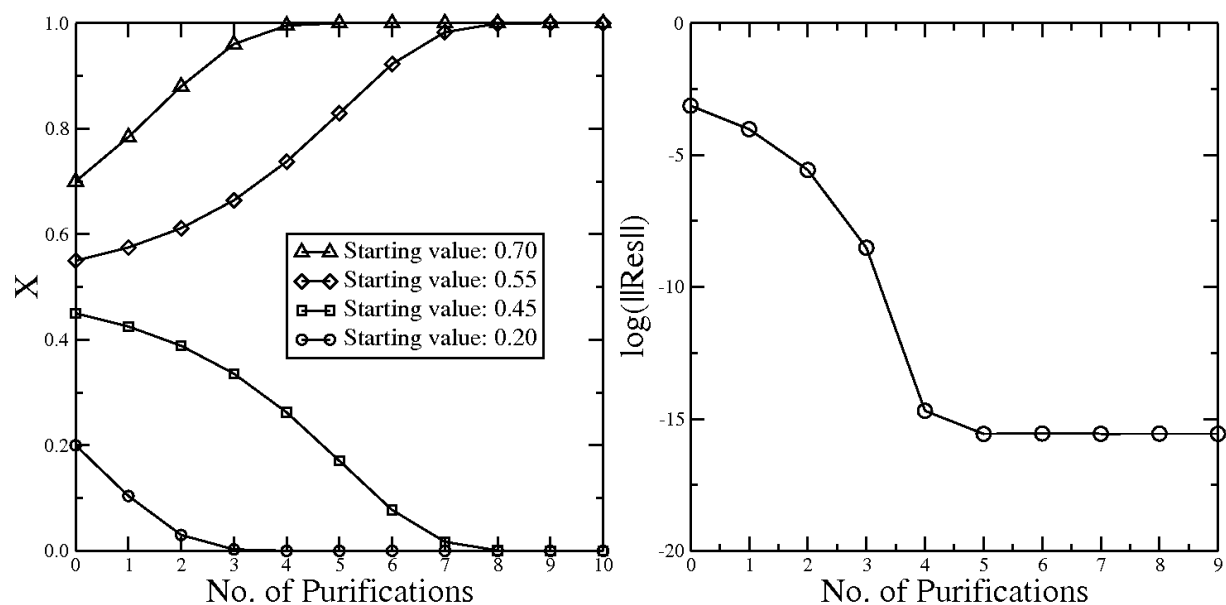


Figure B.2: Localization of electrons at the example of Peierls distortion. The 1-particle reduced density matrix $\rho_1(r; r')$ is plotted for $r = 0$ (center of chain) and different values of r' . The full line indicates the band of hydrogen molecules (molecular) with intra- and intermolecular distances of 1.40 a.u. and 3.32 a.u., respectively. The broken line represents the system of equidistant hydrogen atoms (metallic) with distance 2.36 a.u.. In the lower right corner the MO energies are depicted, both systems contain 142 hydrogen atoms calculated with RHF/6-311G**.

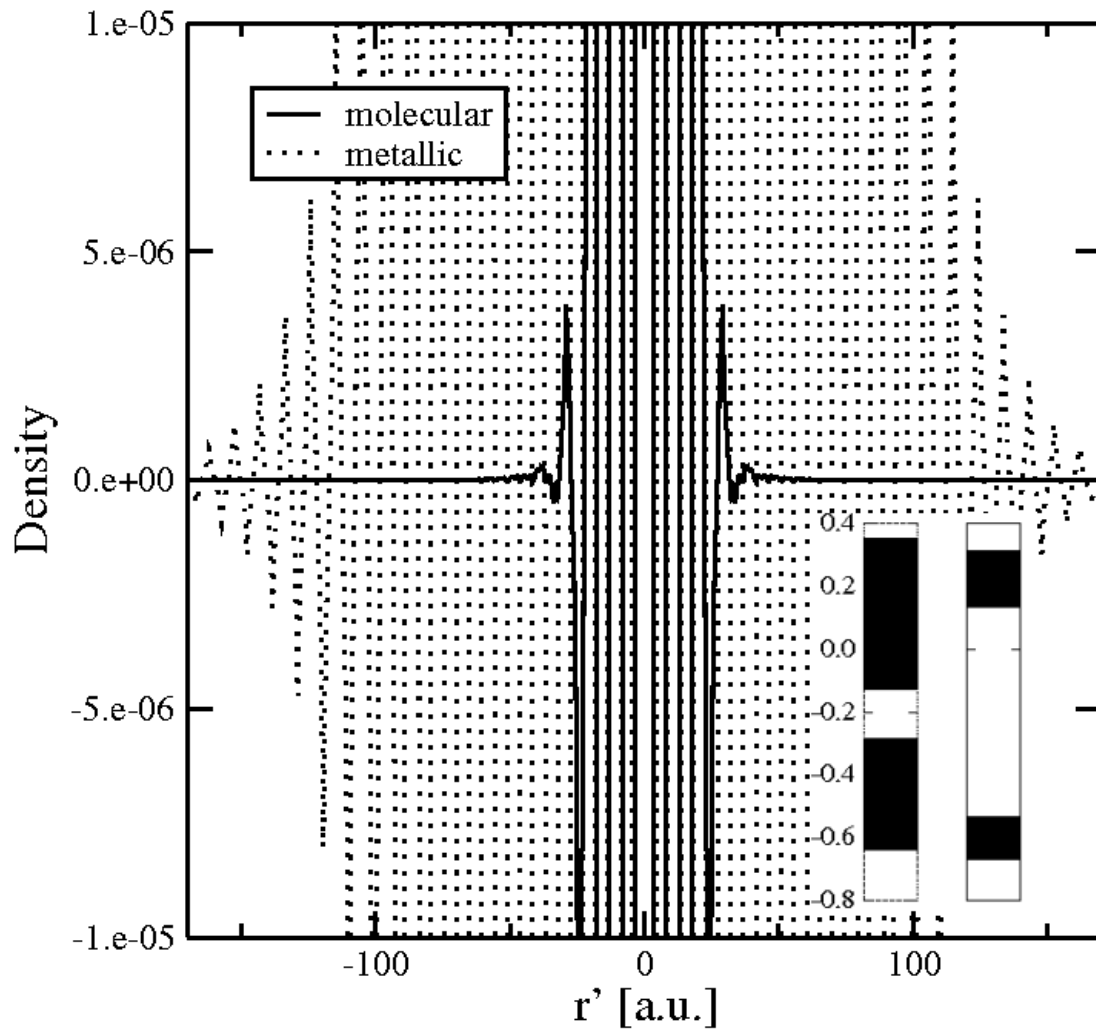


Figure B.3: GIAO-HF/6-31G* of linear alkanes (thr: 10^{-6} , thr_{SA}: 10^{-6}) on Intel Xeon EM64T 3.6GHz processor (Linux): C₂₀H₄₂, C₄₀H₈₂, C₈₀H₁₆₂, C₁₆₀H₃₂₂.

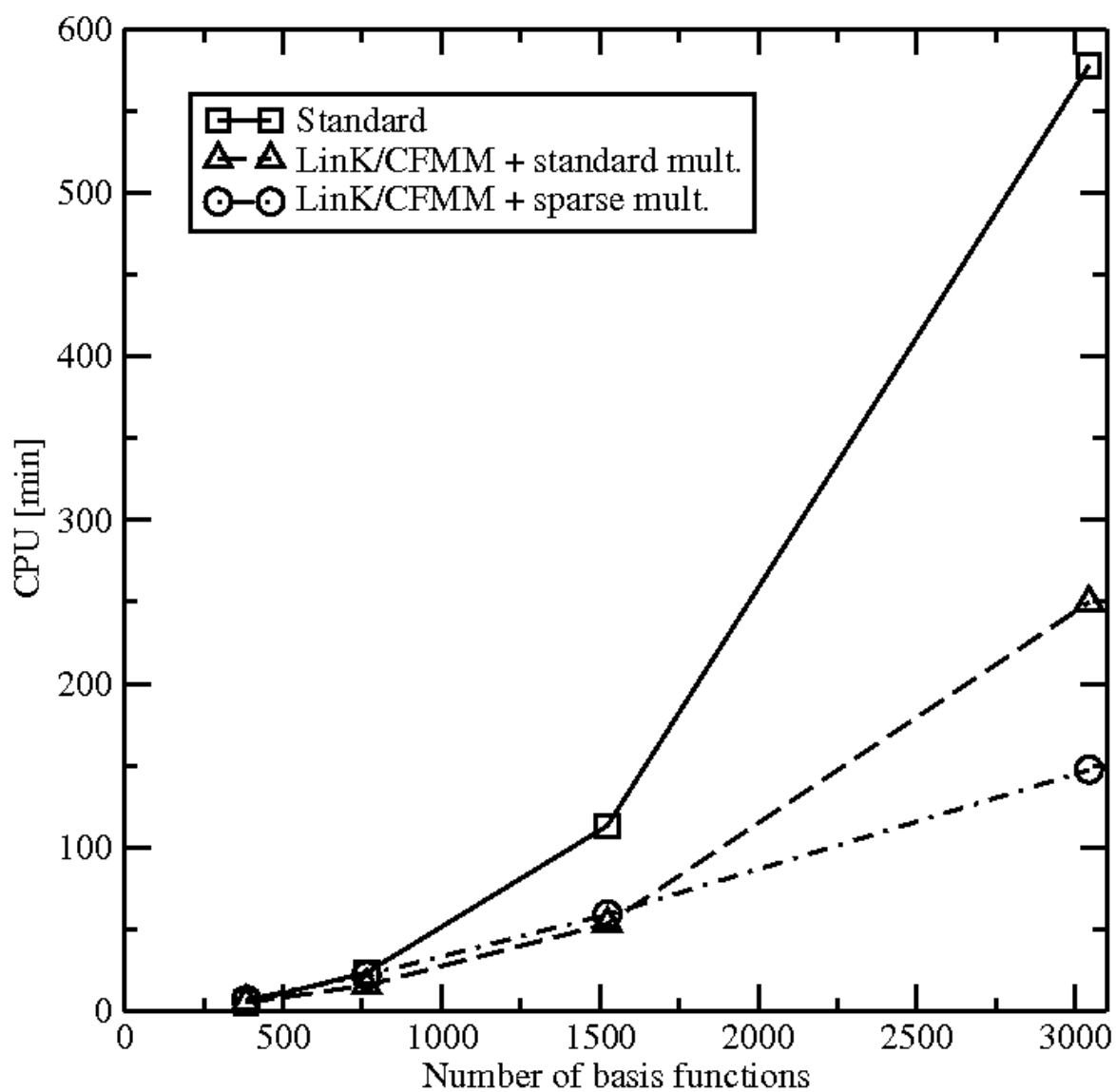


Figure B.4: GIAO-KS-DFT BP86(VWN5)/6-31G* of 1-4 connected amylose chains (thr: 10^{-7} , thr_{SA}: 10^{-7}) on Intel Xeon EM64T 3.6GHz (Linux): α -D-glucose, 2 x α -D-glucose, 4 x α -D-glucose, 8 x α -D-glucose, 16 x α -D-glucose, 32 x α -D-glucose.

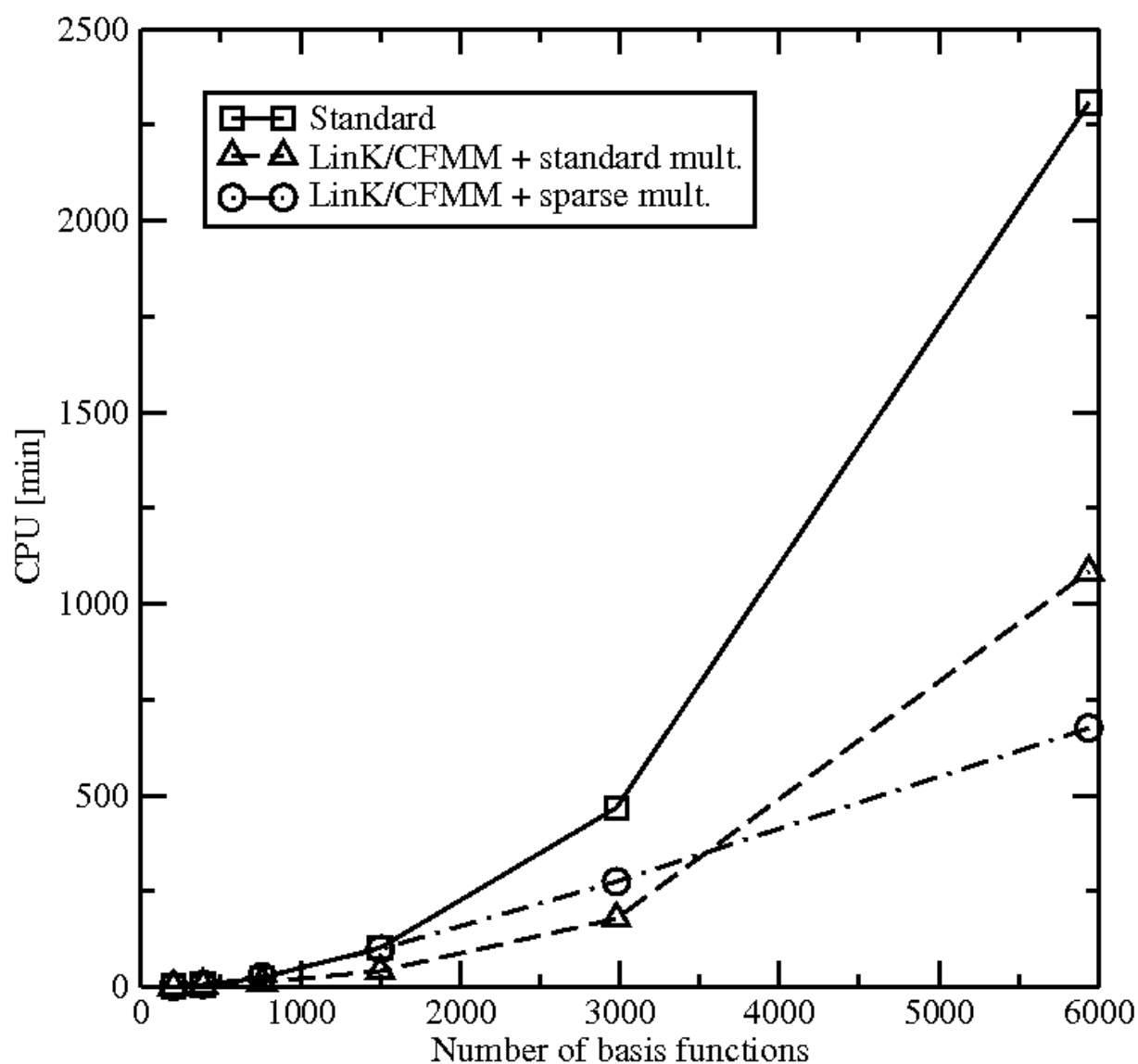


Figure B.5: Scheme of D-CPSCF algorithm.

```

Loop over batches of perturbations
  Form RHS
  Initial Loop Level-1
    Conjugate Gradient to solve  $\mathbf{A}_1\mathbf{x} = \mathbf{b}$ 
  End of Loop
  Loop Level-2:
    Form  $\mathbf{A}_2\mathbf{x}$ 
    DIIS: extrapolation of  $\mathbf{A}_2\mathbf{x}, \mathbf{P}^x, \text{Res}_2 = \|\mathbf{b} - [\mathbf{A}_2\mathbf{x} + \mathbf{A}_1\mathbf{x}]\|$  and
       $\text{Res}_1 = \|\mathbf{b} - \mathbf{A}_1\mathbf{x}\|$ 
    If ( $\|\text{Res}_2\| < \text{convcrit}$ ): exit Loop
    Loop Level-1
      Conjugate gradient to solve  $\mathbf{A}_1\mathbf{x} = \mathbf{b} - \mathbf{A}_2\mathbf{x}$ 
    End of Loop
  End of Loop
End of Loop

```

Figure B.6: TDDFT-BP86(VWN)/6-31G* to calculate the static polarizability tensor of linear alkanes (thr: 10^{-7} , thr_{SA}: 10^{-7}) on Intel Xeon EM64T 3.6GHz (Linux): C₂₀H₄₂, C₄₀H₈₂, C₈₀H₁₆₂, C₁₆₀H₃₂₂, C₂₄₀H₄₈₂.

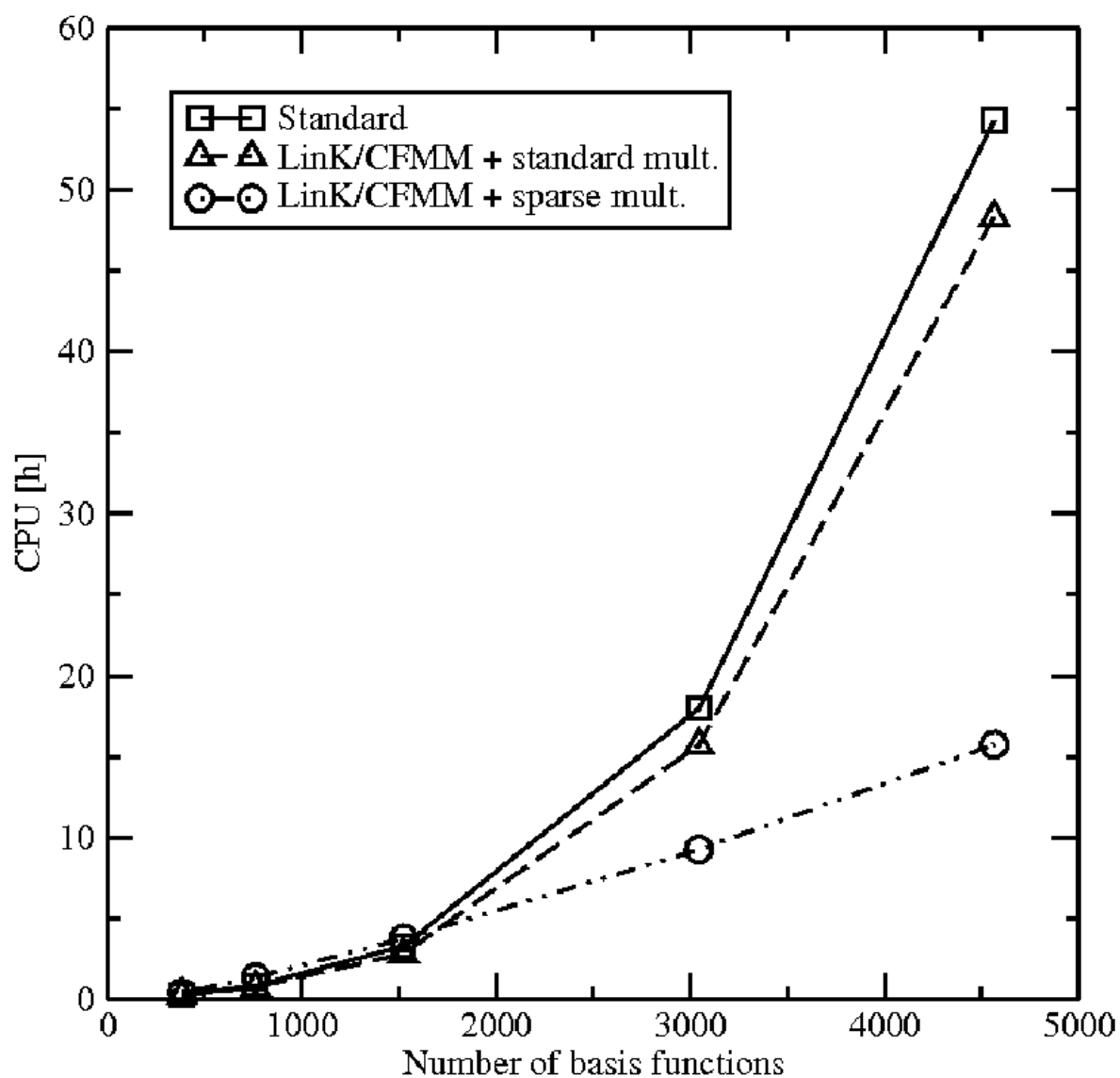


Figure B.7: Frequency-dependent polarizability for hydrogen fluoride with BP86(VWN)/6-31G*. Shaded areas in the left figure sign regions close to the poles of the polarization propagator; on the right the dots centered in a shaded area denote frequencies that equal virtual-occupied orbital differences.

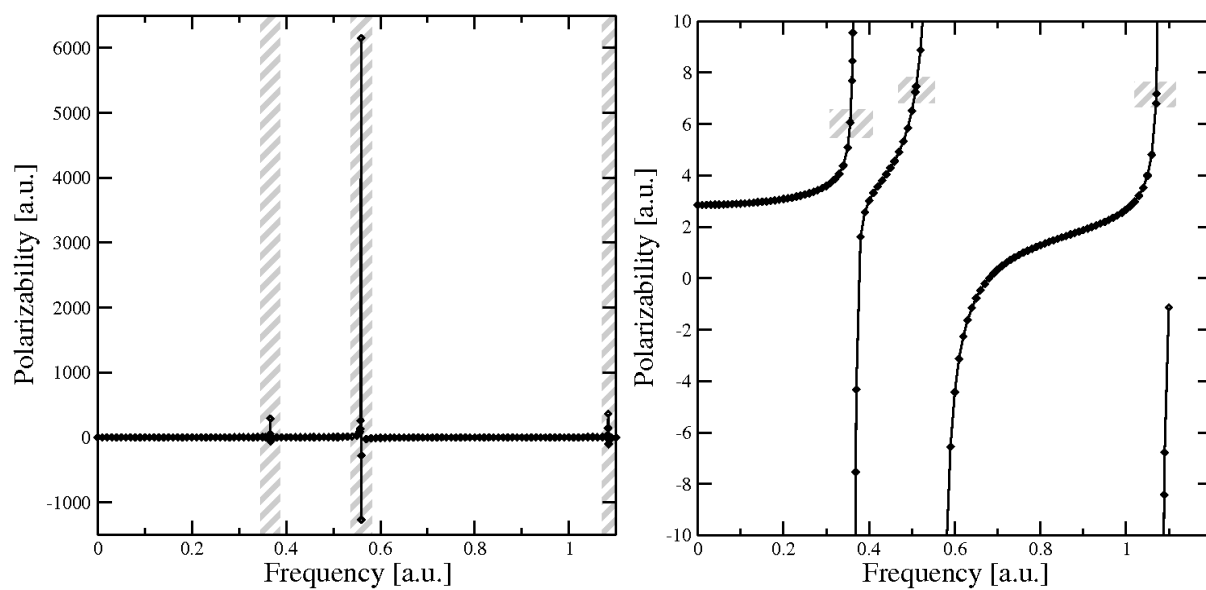


Figure B.8: TDDFT-BP86(VWN)/6-31G* using Wigner's $(2n + 1)$ rule to calculate the static hyperpolarizability tensor of linear alkanes (thr: 10^{-7} , thr_{SA}: 10^{-7}) on Intel Xeon EM64T 3.6GHz (Linux): C₂₀H₄₂, C₄₀H₈₂, C₈₀H₁₆₂, C₁₆₀H₃₂₂, C₂₄₀H₄₈₂. Only the times for the calculation of $\mathbf{V}_{xc}^{(3)}(\mathbf{P}^x, \mathbf{P}^y)$ and the evaluation of eq. 3.191 are given, the first order quantities are provided by a first order TDDFT calculation (see Fig. B.6).

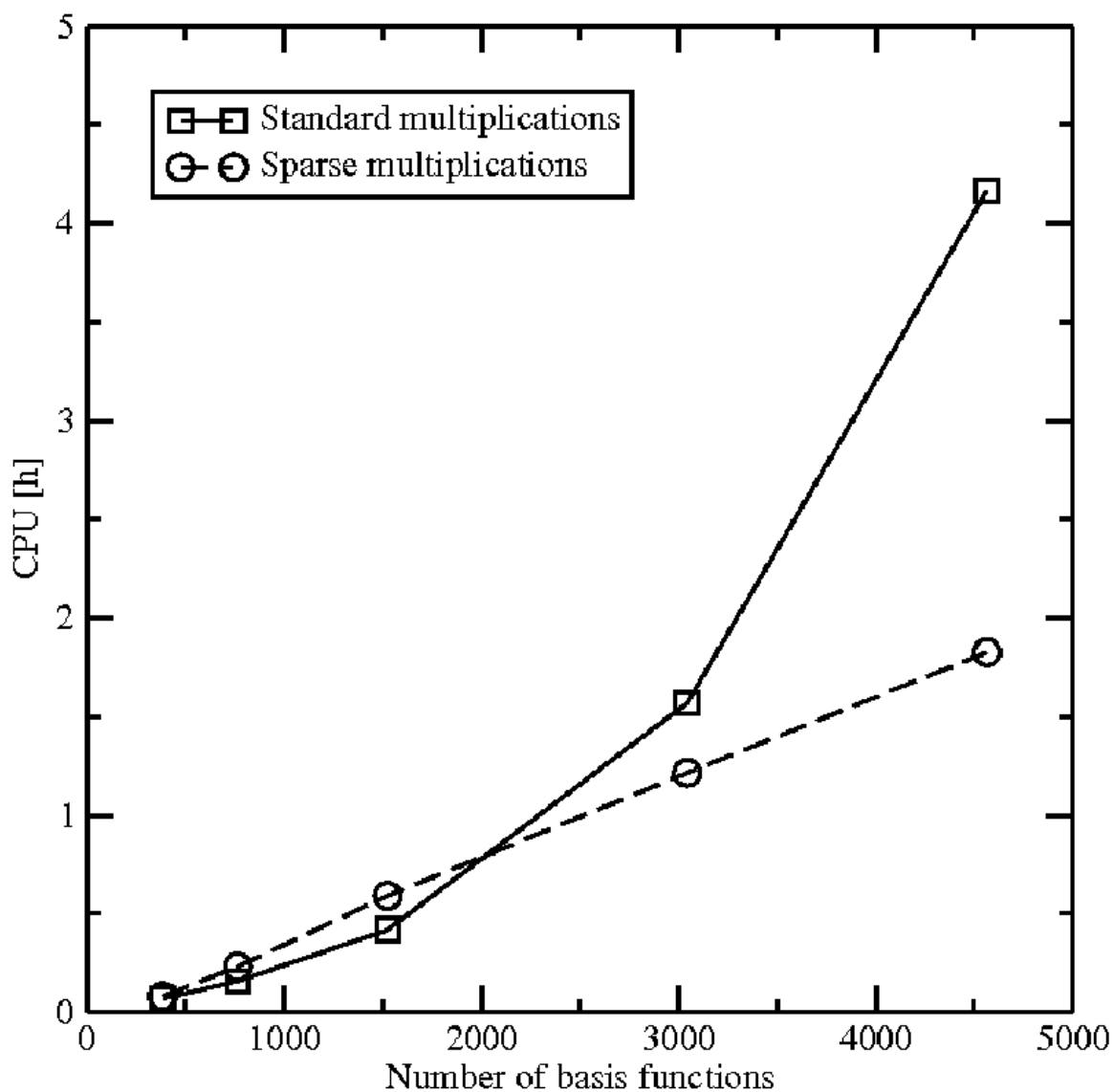


Figure B.9: HCF₃ molecule with HF/6-31G*. Fluctuations in the maximal element of \mathbf{P} within 1000 purification steps for different screening schemes. $\tilde{\mathbf{P}} = 3\mathbf{PSP} - 2\mathbf{PSPSP}$ (straight line: not screened, broken lines: screened).

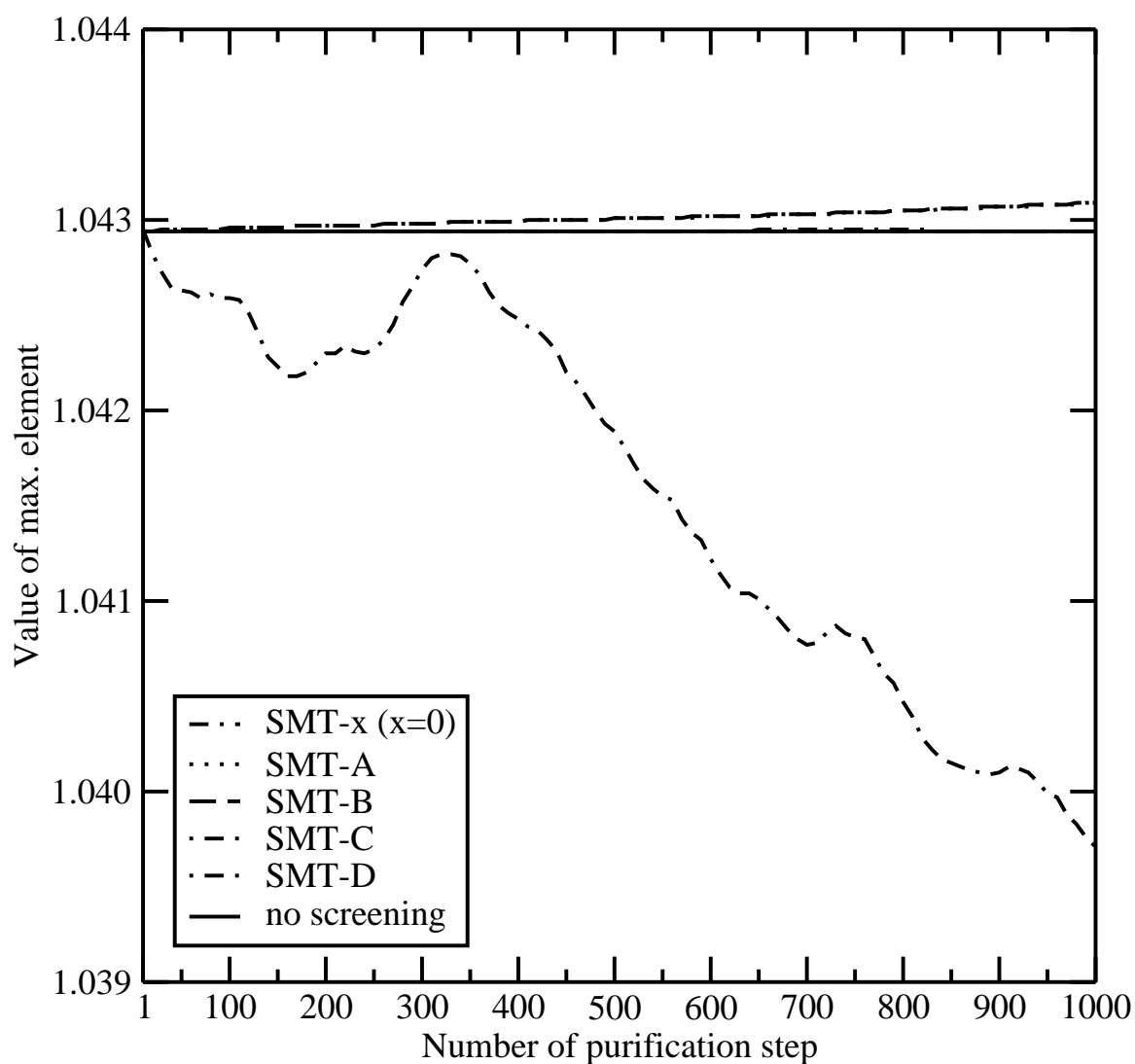


Figure B.10: HCF₃ molecule with HF/6-31G*. Flop-counting for 1000 purification steps with $\text{thr}_{\text{SA}} = 10^{-5}$ for different screening schemes.

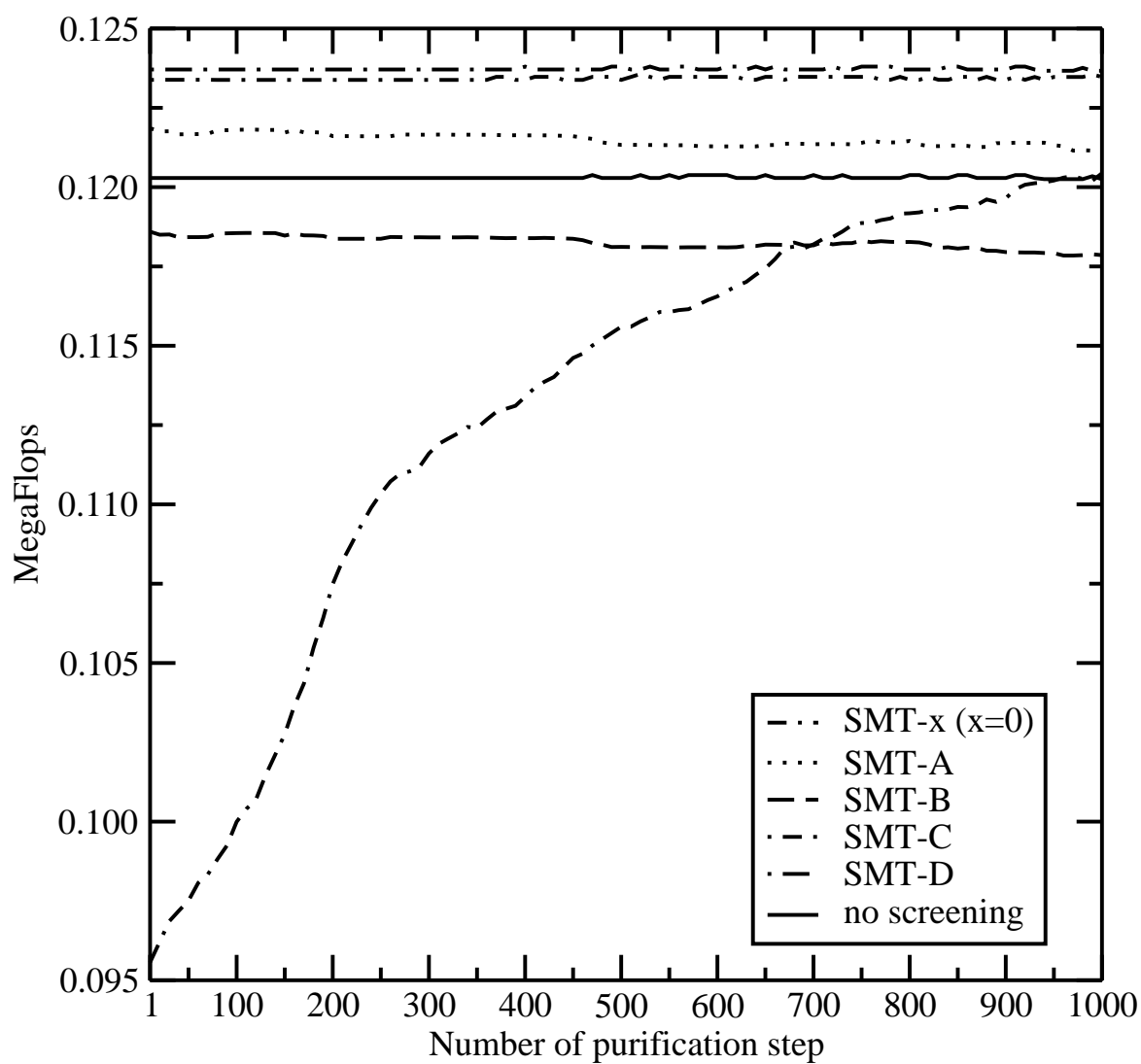


Figure B.11: HCF₃ molecule with HF/6-31G*. Sparsity in percent of **P** within 1000 purification steps with $\text{thr}_{\text{SA}} = 10^{-5}$ for different screening schemes.

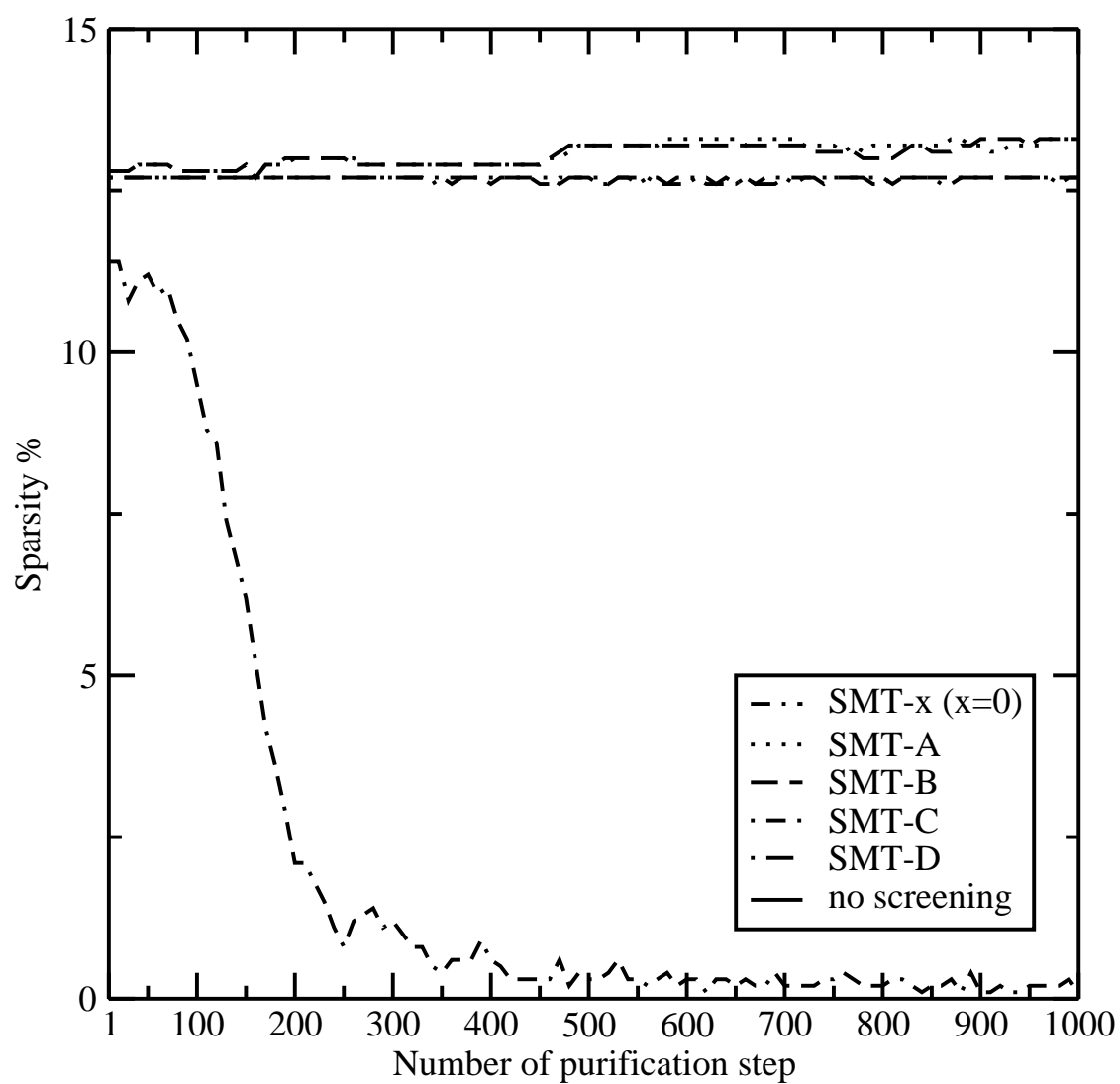


Figure B.12: Amylose chain of 8 a-D-glucose units. Sum of Giga-Flops for different Level-1 iterations within GIAO-KS-DFT BP86(VWN)/6-31G* ($\text{thr}_{\text{SA}} = 10^{-7}$).

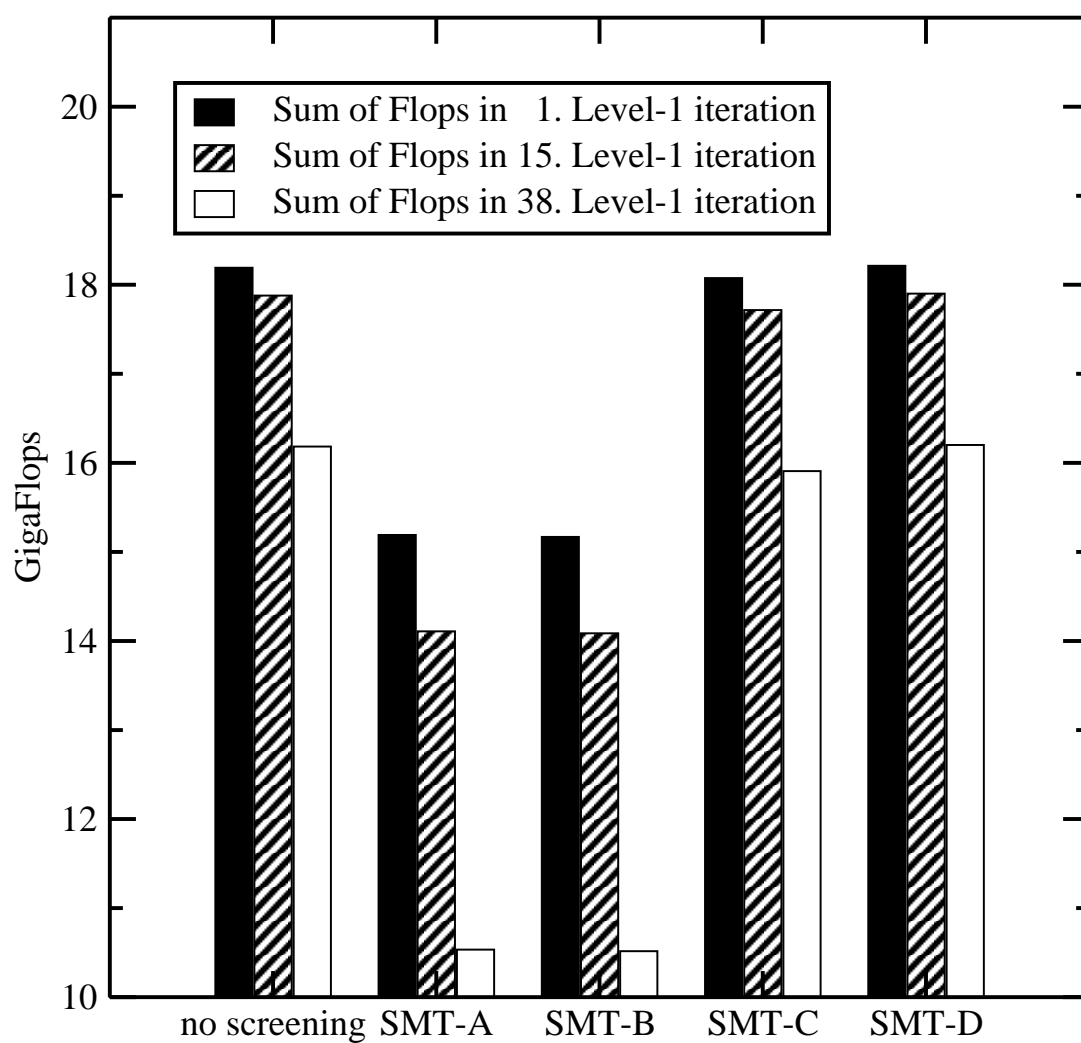


Figure B.13: Applying screening techniques to sparse multiplications. Calculations as in Fig. B.4, times for linear equation solver without screening and SMT-A.

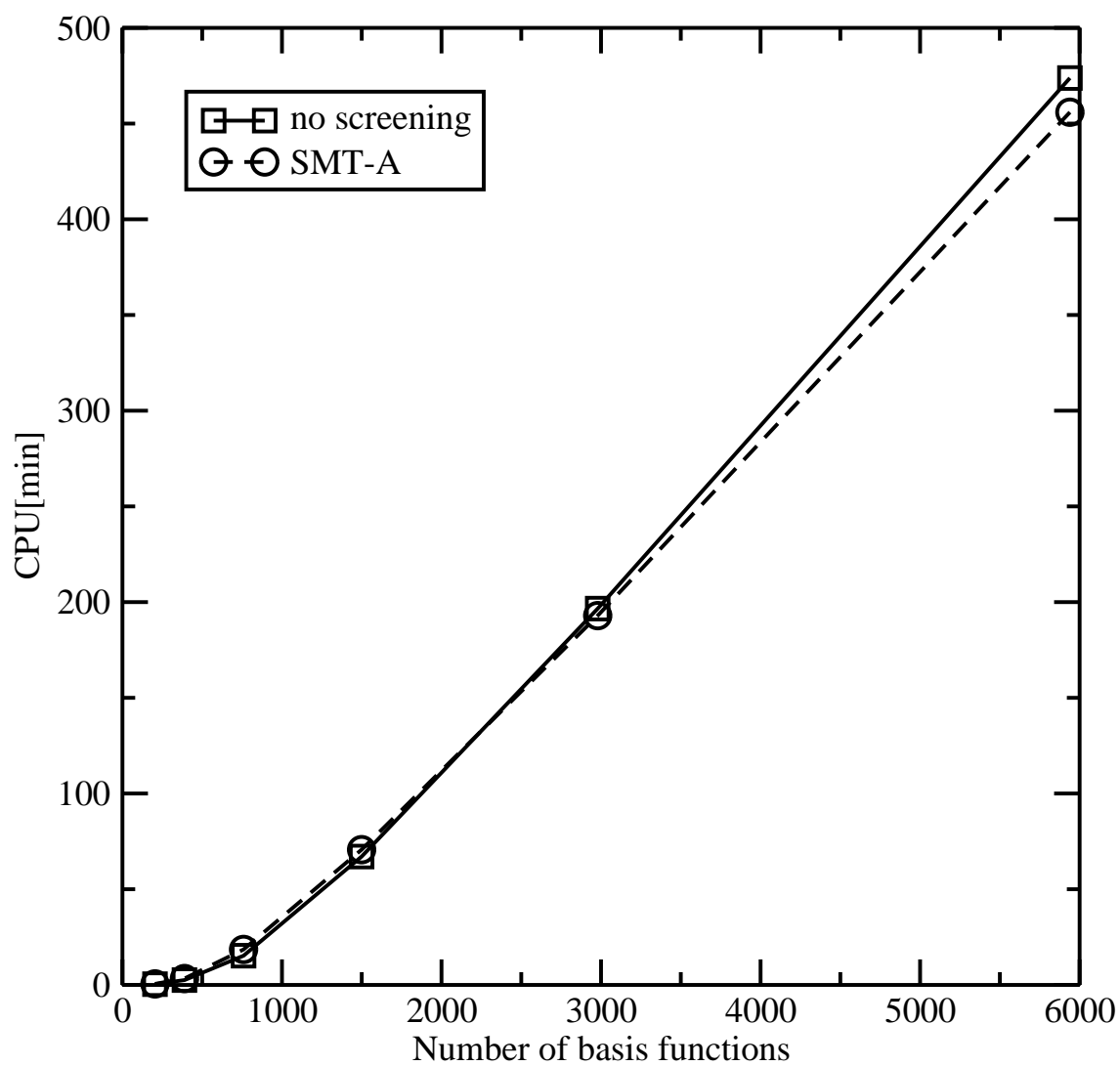


Figure B.14: Scheme of SMT-A/B/C/x screening modifications of sparse matrix multiplication routine. n_A is the real length of the actual row of matrix \mathbf{A} , n_{AB1} the minimum of n_A and the maximum of real lengths of the columns of the corresponding elements in the actual row of matrix \mathbf{B} , n_{AB2} the minimum of n_A and the maximum of real lengths of all columns of matrix \mathbf{B} , n_{AB3} the minimum of n_A and the average real length of the columns of matrix \mathbf{B} .

Preordering of elements of each row of matrix \mathbf{B} in decreasing order

Loop over N rows of \mathbf{A} : $\sum_{i=1}^N$

Loop over constant number of elements in row i : $\sum_{k=1}^{const} A_{ik}$

If SMT-A: $\text{thr}_{\text{eff}} = \text{thr}_{\text{SA}}/n_{AB1}$

If SMT-B: $\text{thr}_{\text{eff}} = \text{thr}_{\text{SA}}/n_{AB2}$

If SMT-C: $\text{thr}_{\text{eff}} = \text{thr}_{\text{SA}}/n_{AB3}$

If SMT-D: $\text{thr}_{\text{eff}} = \text{thr}_{\text{SA}}/n_A$

If SMT-x: $\text{thr}_{\text{eff}} = \text{thr}_{\text{SA}} \cdot 10^{-x}$

Determine contraction length N_{SMT} with respect to $\text{thr}_{\text{eff}}/\text{abs}(A_{ik})$

Loop over constant number of elements in row k of matrix \mathbf{B} : $\sum_{j=1}^{N_{\text{SMT}}} B_{kj}$

Add product to corresponding element in matrix \mathbf{C} : $C_{ij} = C_{ij} + A_{ik}B_{kj}$

End of Loop

End of Loop

Store significant elements of row i of matrix \mathbf{C}

End of Loop

Figure B.15: Plot of the first (left) and second (right) derivative of the 1s-function of lithium in the cc-pVTZ basis as function of the electron-nuclear distance r . The broken line is the derivative of the simple contracted Gaussian while the full line shows the basis function that has been fitted to a Slater-type function in the near-nucleus region.

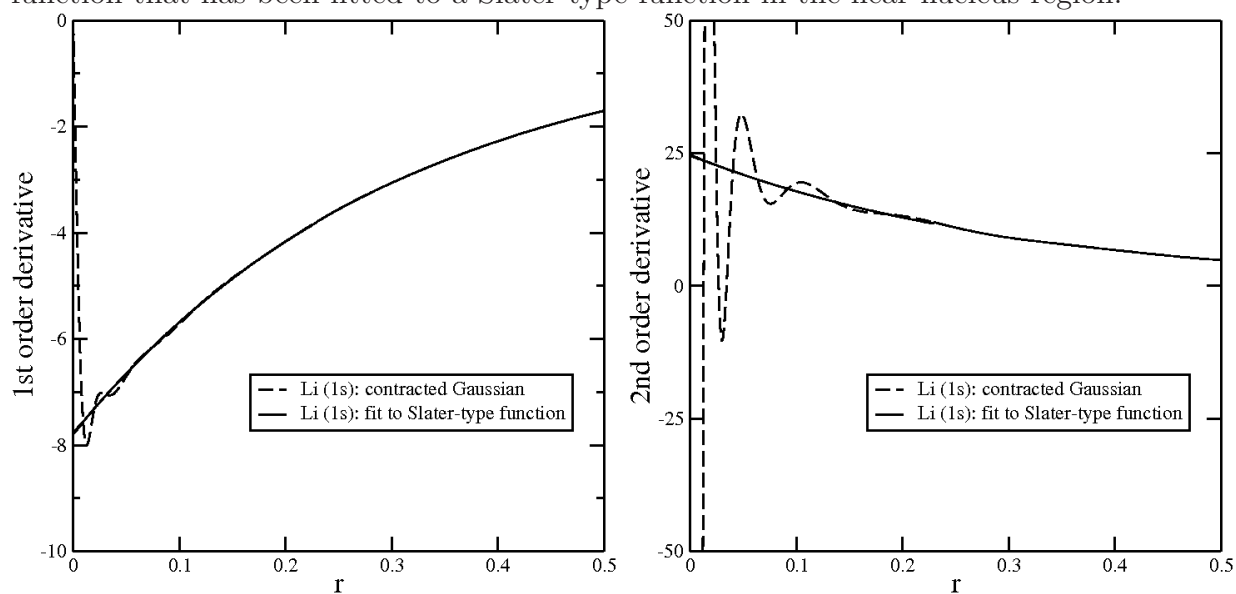


Figure B.16: Scheme of algorithm to compute ρ_N and its 1st and 2nd derivatives with respect to the single electron positions within N -PDM VQMC.

For α - and β -electrons:

- Build χ , $\chi^{xyz,1st}$, χ^{2nd} into sparse matrix (5 x $\mathcal{O}(N)$)
- Multiplies: (6 x $\mathcal{O}(N)$)

$$\begin{aligned} \mathbf{D}_\rho &= \chi \mathbf{P} \chi^\dagger \\ \mathbf{D}_\rho^{u,1st} &= \chi^{u,1st} \mathbf{P} \chi^\dagger \\ \mathbf{D}_\rho^{2nd} &= \chi^{2nd} \mathbf{P} \chi^\dagger \end{aligned}$$

- Cholesky decomposition of \mathbf{D}_ρ : (4 x $\mathcal{O}(N)$)
 - Build Cholesky factor: L
 - Form determinant: $\det \mathbf{D}_\rho = \rho_N = \prod_i L_{ii}$
 - Build inverse Cholesky factor: L^{-1}
 - Build inverse: $\tilde{\mathbf{D}}_\rho = (L^{-1})^\top L^{-1}$
- Determine values of gradients of Laplacian: (4 x $\mathcal{O}(N)$)

$$\begin{aligned} \rho_N(\mathbf{R}; \mathbf{R}')^{-1} \nabla_{u,i} \rho_N(\mathbf{R}; \mathbf{R}') &= \tilde{D}_{\rho,ij} D_{\rho,ij}^{u,1st} \\ \rho_N(\mathbf{R}; \mathbf{R}')^{-1} \nabla_i^2 \rho_N(\mathbf{R}; \mathbf{R}') &= \tilde{D}_{\rho,ij} D_{\rho,ij}^{2nd} \end{aligned}$$

Figure B.17: N -PDM VQMC calculations of a series of linear alkanes (basis: cc-pVTZ). CPU times needed to evaluate the basis functions of a single random walker an 1000 sampling steps. The three numbers that are given for the calculations with screening and sparse algebra denote the threshold settings. E.g. $x/y/z$: 10^{-x} for basis function screening, 10^{-y} as general sparse matrix cut-off and 10^{-z} as compression threshold for the discrete Fock-Dirac density P .

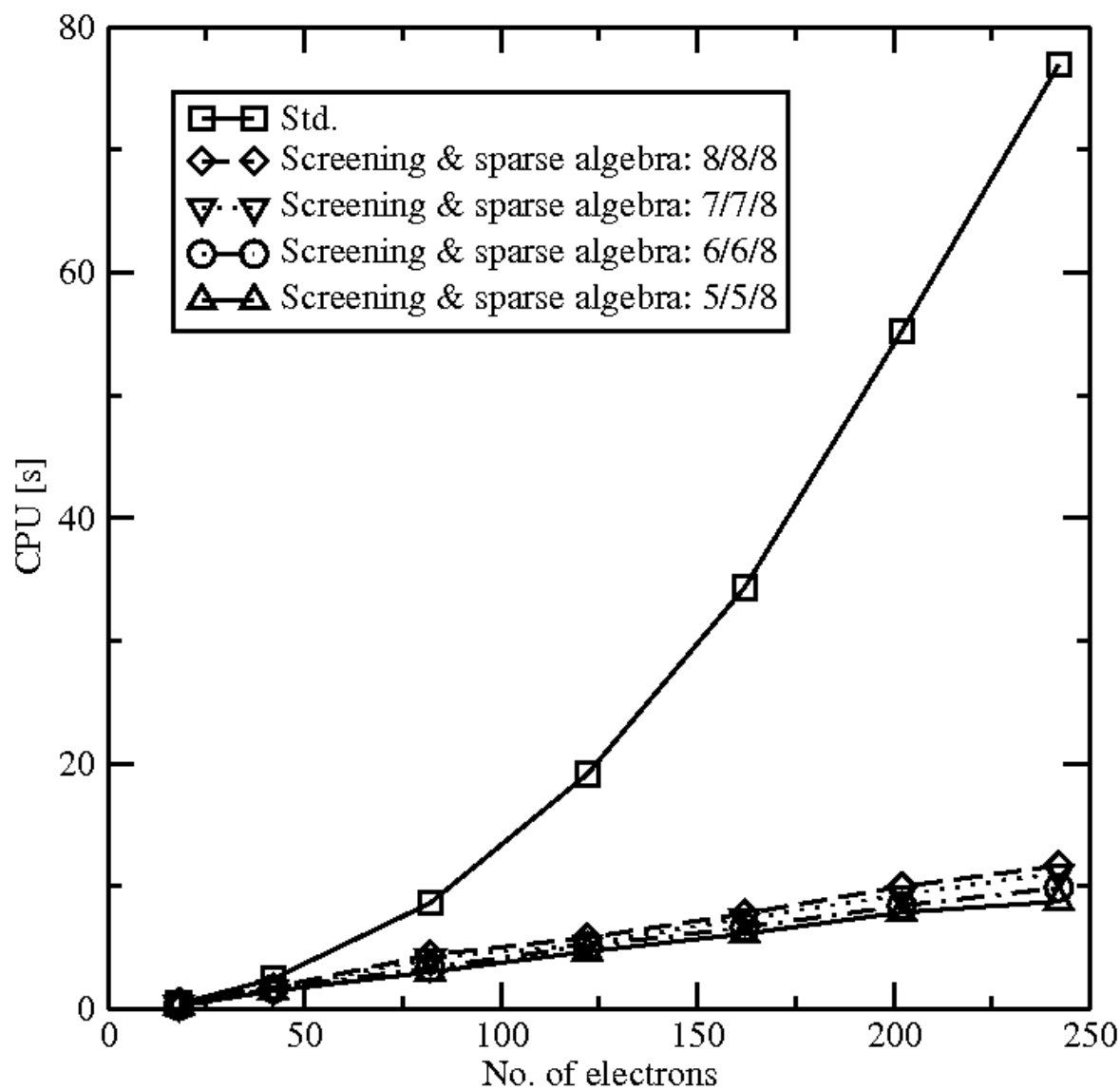


Figure B.18: N -PDM VQMC calculations of a series of linear alkanes (basis: cc-pVTZ). CPU times needed for the matrix multiplies of a single random walker an 1000 sampling steps. The three numbers that are given for the calculations with screening and sparse algebra denote the threshold settings. E.g. $x/y/z$: 10^{-x} for basis function screening, 10^{-y} as general sparse matrix cut-off and 10^{-z} as compression threshold for the discrete Fock-Dirac density P .

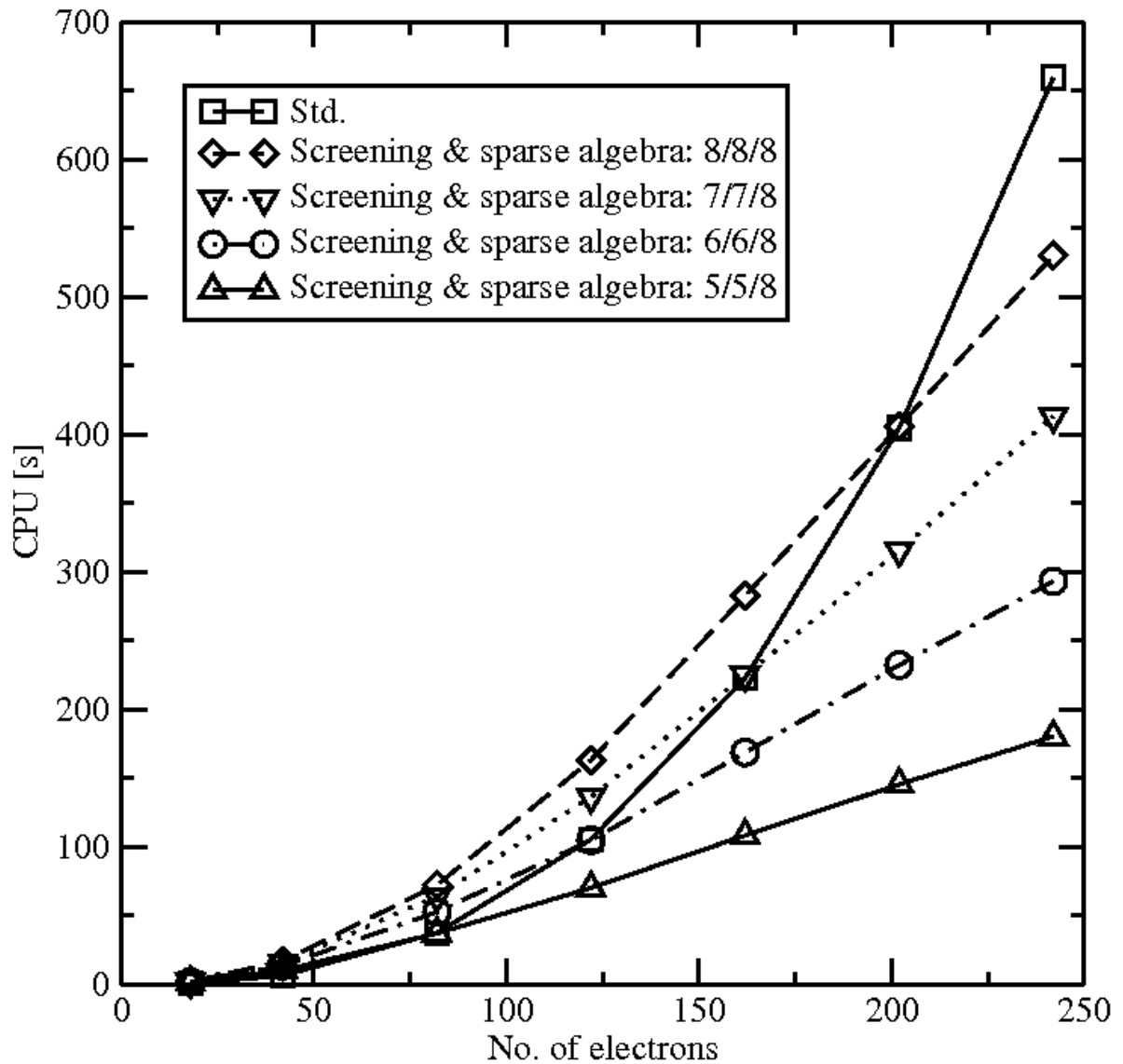


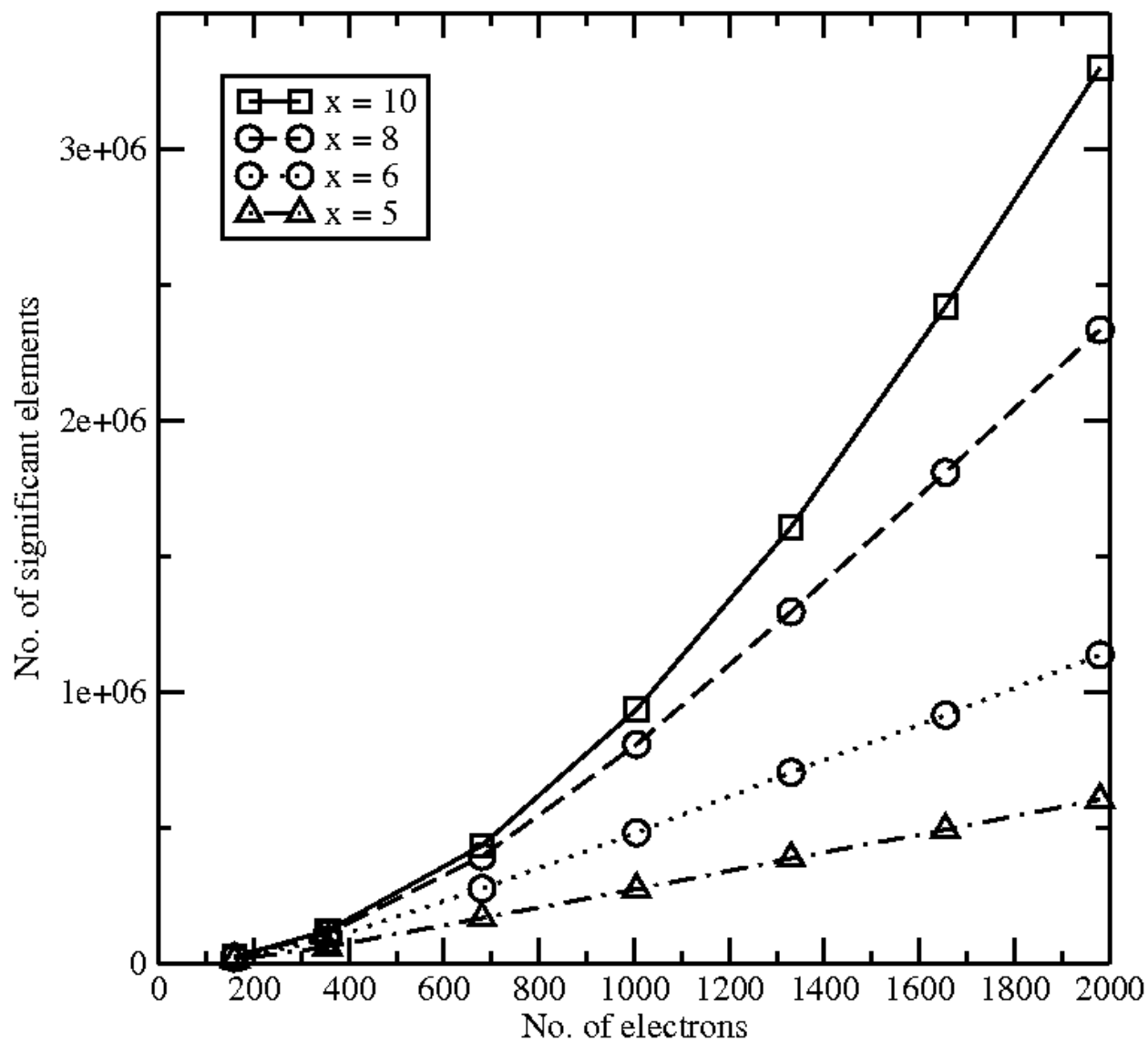
Figure B.19: Number of significant elements in \mathbf{P} for a given threshold 10^{-x} .

Figure B.20: Scheme of algorithm to compute ρ_N and its 1st and 2nd derivatives with respect to the single electron positions within N -PDM FN-DQMC.

For α - and β -electrons:

- Build $\chi(\mathbf{R}_\alpha)$, $\chi^{xyz,1st}(\mathbf{R}_\alpha)$, $\chi^{2nd}(\mathbf{R}_\alpha)$ into sparse matrix ($5 \times \mathcal{O}(N)$)
- Multiplies: ($6 \times \mathcal{O}(N)$)

$$\mathbf{P}(\chi(\mathbf{R}_\alpha))^\dagger \quad (\text{save matrix for the next cycle})$$

$$\mathbf{D}_\rho = \chi(\mathbf{R}_\alpha)\mathbf{P}(\chi(\mathbf{R}_{\alpha-1}))^\dagger$$

$$\mathbf{D}_\rho^{u,1st} = \chi^{u,1st}(\mathbf{R}_\alpha)\mathbf{P}(\chi(\mathbf{R}_{\alpha-1}))^\dagger$$

$$\mathbf{D}_\rho^{2nd} = \chi^{2nd}(\mathbf{R}_\alpha)\mathbf{P}(\chi(\mathbf{R}_{\alpha-1}))^\dagger$$

- LU decomposition of \mathbf{D}_ρ : ($4 \times \mathcal{O}(N)$)
 - Build LU factors (UMFPACK): L, U
 - Form determinant: $\det \mathbf{D}_\rho(\mathbf{R}_\alpha; \mathbf{R}_{\alpha-1}) = \rho_N(\mathbf{R}_\alpha; \mathbf{R}_{\alpha-1}) = \prod_i L_{ii}$
 - Build inverse by backsubstitution: $\tilde{D}_\rho(\mathbf{R}_\alpha; \mathbf{R}_{\alpha-1})$
 - Transpose inverse: $(\tilde{D}_\rho(\mathbf{R}_\alpha; \mathbf{R}_{\alpha-1}))^\dagger$
- Determine values of gradients of Laplacian: ($4 \times \mathcal{O}(N)$)

$$\rho_N(\mathbf{R}_\alpha; \mathbf{R}_{\alpha-1})^{-1} \nabla_{u,i} \rho_N(\mathbf{R}_\alpha; \mathbf{R}_{\alpha-1}) = (\tilde{D}_\rho(\mathbf{R}_\alpha; \mathbf{R}_{\alpha-1}))_{ij}^\dagger D_\rho^{u,1st}(\mathbf{R}_\alpha; \mathbf{R}_{\alpha-1})_{ij}$$

$$\rho_N(\mathbf{R}_\alpha; \mathbf{R}_{\alpha-1})^{-1} \nabla_i^2 \rho_N(\mathbf{R}_\alpha; \mathbf{R}_{\alpha-1}) = (\tilde{D}_\rho(\mathbf{R}_\alpha; \mathbf{R}_{\alpha-1}))_{ij}^\dagger D_\rho^{2nd}(\mathbf{R}_\alpha; \mathbf{R}_{\alpha-1})_{ij}$$

Figure B.21: N -PDM FN-DQMC calculations of a series of linear alkanes (basis: cc-pVTZ). CPU times for for 1000 sampling steps of a single random walker. The LU factorization and the construction of the inverse determinant matrix $\tilde{\mathbf{D}}_\rho$ is performed with the routines of the UMFPACK library [165], before the LU factorization as in Fig. B.22. The three numbers that are given for the calculations with screening and sparse algebra denote the threshold settings. E.g. $x/y/z$: 10^{-x} for basis function screening, 10^{-y} as general sparse matrix cut-off and 10^{-z} as compression threshold for the discrete Fock-Dirac density \mathbf{P} .

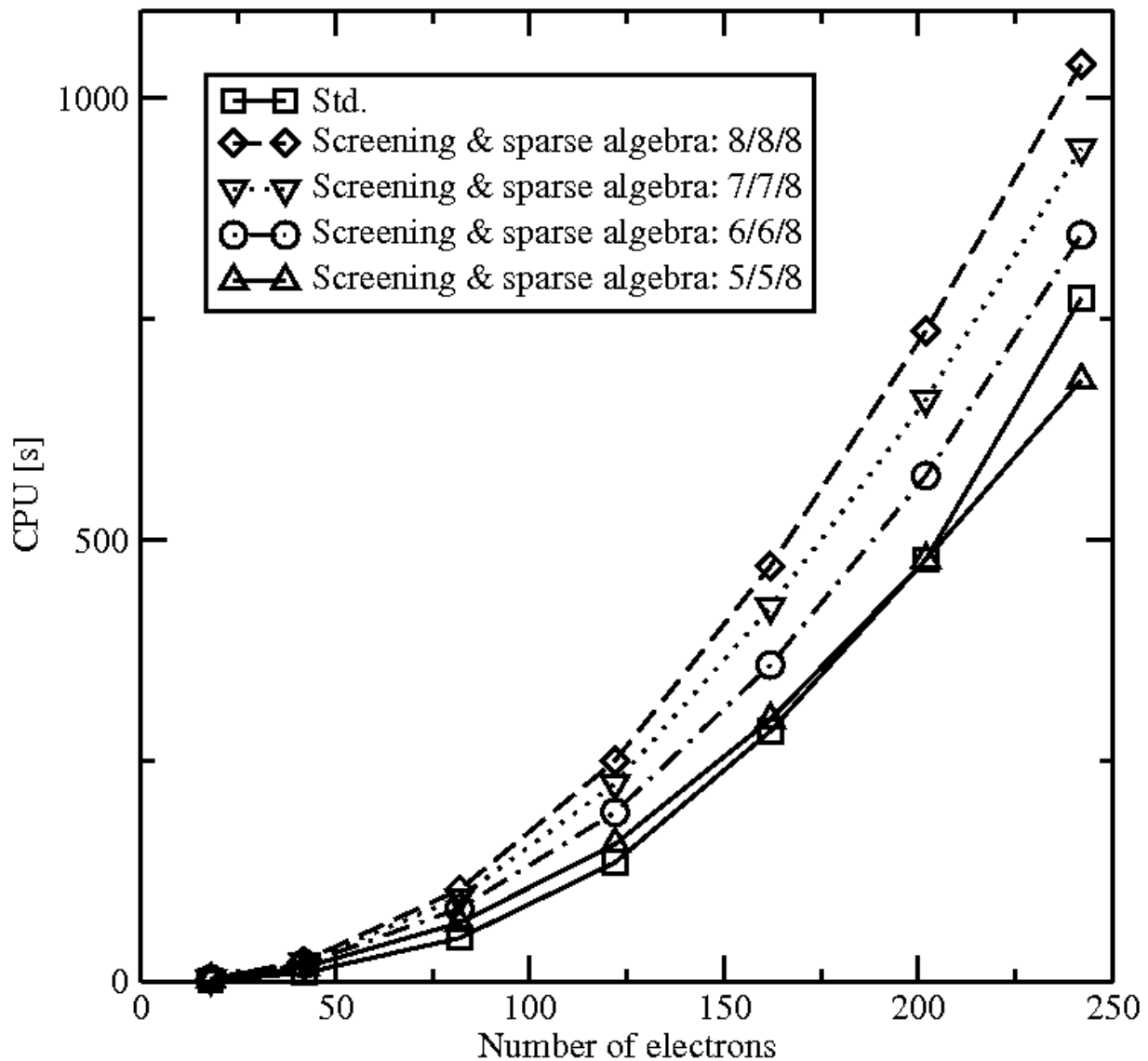
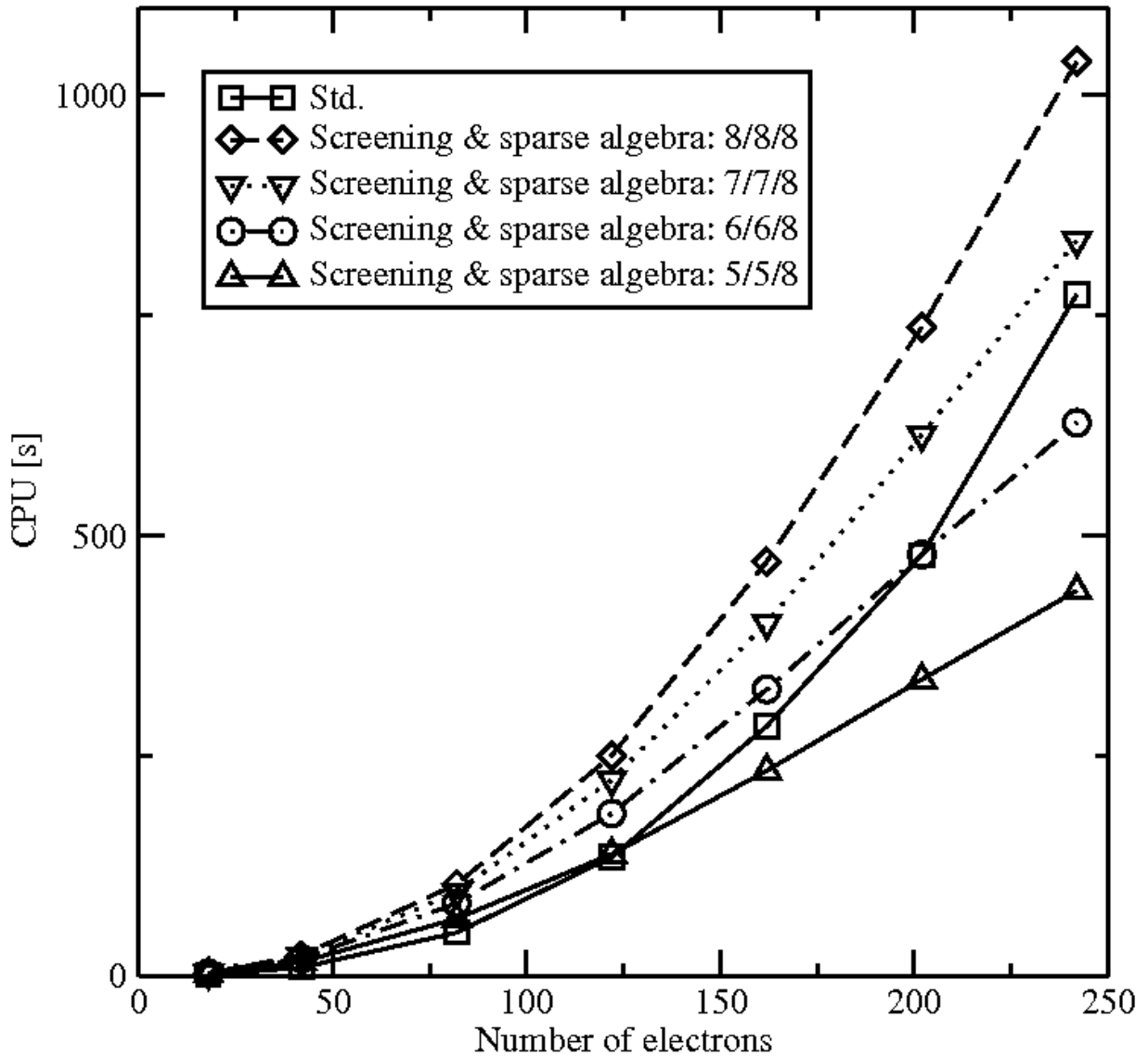


Figure B.22: N -PDM FN-DQMC calculations of a series of linear alkanes (basis: cc-pVTZ). CPU times for for 1000 sampling steps of a single random walker. The determinant matrix is recompressed with respect to $\text{thr}_{\mathbf{P}}$. The routines of the UMFPACK library are used to perform the LU factorization and the construction of the inverse determinant matrix $\tilde{\mathbf{D}}_{\rho}$. The three numbers that are given for the calculations with screening and sparse algebra denote the threshold settings. E.g. $x/y/z$: 10^{-x} for basis function screening, 10^{-y} as general sparse matrix cut-off and 10^{-z} as compression threshold for the discrete Fock-Dirac density \mathbf{P} .



Appendix C

Tables

Table C.1: Structure of para-nitroaniline as Cartesian coordinates in Å (C_{2v} symmetry). The geometry is obtained from a geometry optimization with B3LYP/6-31G*

Atom	x	y	z
C	0.000000	0.000000	2.083262
N	0.000000	0.000000	3.450969
H	0.861349	0.000000	3.972676
C	-1.215837	0.000000	1.362624
H	-2.158663	0.000000	1.903552
C	-1.215378	0.000000	-0.021632
H	-2.141991	0.000000	-0.582542
C	0.000000	0.000000	-0.711767
N	0.000000	0.000000	-2.165325
O	-1.090947	0.000000	-2.744236
O	1.090947	0.000000	-2.744236
H	-0.861349	0.000000	3.972676
C	1.215378	0.000000	-0.021632
H	2.141991	0.000000	-0.582542
C	1.215837	0.000000	1.362624
H	2.158663	0.000000	1.903552

Table C.2: Frequency-dependent polarizabilities of para-nitroaniline with HF/6-31G*. Only non-zero elements of the tensor are given in 10^{-25} esu (1 a.u. = $1.48 \times 10^{-25} \text{cm}^{-3}$).

Method	Frequency [a.u.]/[eV]	α_{xx}	α_{yy}	α_{zz}	$\bar{\alpha}$
D-TDSCF	0.0000/0.000	129.61	42.80	181.18	117.87
	0.0240/0.650	130.02	42.83	182.38	118.41
	0.0478/1.300	131.27	42.90	186.17	120.11
MO-TDSCF	0.0000/0.000	129.61	42.80	181.18	117.87
	0.0240/0.650	130.02	42.83	182.38	118.41
	0.0478/1.300	131.27	42.90	186.17	120.11

Table C.3: Frequency-dependent polarizabilities of para-nitroaniline with B3LYP/6-31G*. Only non-zero elements of the tensor are given in 10^{-25} esu (1 a.u. = $1.48 \times 10^{-25} \text{cm}^{-3}$).

Method	Frequency [a.u.]/[eV]	α_{xx}	α_{yy}	α_{zz}	$\bar{\alpha}$
D-TDDFT	0.0000/0.000	132.34	42.93	208.06	127.78
	0.0240/0.650	132.80	42.96	210.30	128.69
	0.0478/1.300	134.21	43.04	217.67	131.64
MO-TDDFT	0.0000/0.000	132.34	42.93	208.06	127.78
	0.0240/0.650	132.80	42.96	210.30	128.69
	0.0478/1.300	134.21	43.04	217.67	131.64

Table C.4: Frequency-dependent first hyperpolarizabilities of para-nitroaniline with HF/6-31G* and frequency $\omega = 0.024$ a.u. (0.650 eV). Only non-zero elements of the tensor are given in 10^3 esu ($1 \text{ a.u.} = 8.64 \times 10^{-33} \text{g}^{-1/2} \text{cm}^{7/2} \text{s}$).

Method	Property	β_{xxz}	β_{yyz}	β_{zxx}	β_{zyy}	β_{zzz}	$\bar{\beta}_z$
D-TDSCF	$\beta(0; 0, 0)$	-0.848	-0.002	-0.848	-0.002	4.652	3.802
	$\beta(-2\omega, +\omega, +\omega)$	-0.877	-0.002	-0.902	-0.002	5.038	4.150
	$\beta(0, +\omega, -\omega)$	-0.863	-0.002	-0.855	-0.002	4.775	3.914
MO-TDSCF	$\beta(0; 0, 0)$	-0.848	-0.002	-0.848	-0.002	4.652	3.802
	$\beta(-2\omega, +\omega, +\omega)$	-0.877	-0.002	-0.902	-0.002	5.038	4.150
	$\beta(0, +\omega, -\omega)$	-0.863	-0.002	-0.855	-0.002	4.775	3.914

Table C.5: Frequency-dependent first hyperpolarizabilities of para-nitroaniline with B3LYP/6-31G* and frequency $\omega = 0.024$ a.u. (0.650 eV). Only non-zero elements of the tensor are given in 10^3 esu ($1 \text{ a.u.} = 8.64 \times 10^{-33} \text{g}^{-1/2} \text{cm}^{7/2} \text{s}$).

Method	Property	β_{xxz}	β_{yyz}	β_{zxx}	β_{zyy}	β_{zzz}	$\bar{\beta}_z$
D-TDSCF	$\beta(0; 0, 0)$	-0.658	-0.009	-0.658	-0.009	6.475	5.808
	$\beta(-2\omega, +\omega, +\omega)$	-0.679	-0.010	-0.722	-0.010	7.481	6.778
	$\beta(0, +\omega, -\omega)$	-0.674	-0.009	-0.661	-0.009	6.785	6.106
MO-TDSCF	$\beta(0; 0, 0)$	-0.658	-0.009	-0.658	-0.009	6.475	5.808
	$\beta(-2\omega, +\omega, +\omega)$	-0.679	-0.010	-0.722	-0.010	7.481	6.778
	$\beta(0, +\omega, -\omega)$	-0.674	-0.009	-0.661	-0.009	6.785	6.106

Table C.6: Second harmonic generation of para-nitroaniline with D-TDDFT B3LYP/6-31G*. Only non-zero elements of the tensor are given in 10^3 esu (1 a.u. = $8.64 \times 10^{-33} \text{g}^{-1/2} \text{cm}^7/2\text{s}$).

Frequency [eV]	β_{xxz}	β_{yyz}	β_{zxx}	β_{zyy}	β_{zzz}	$\bar{\beta}_z$
1.170	-0.731	-0.011	-0.942	-0.015	10.928	10.114
1.361	-0.762	-0.012	-1.127	-0.018	13.855	12.955
1.494	-0.788	-0.013	-1.342	-0.022	17.235	16.246

Table C.7: HCF₃ molecule with HF/6-31G* ($N_{\text{BF}}=62$, $N_{\text{BF}}^2=3844$). Explicit sparsity pattern of \mathbf{P} after 1000 purification steps.

Global Threshold	Threshold									
	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
10^{-3} (not screened)	98.0	75.7	49.0	48.0	48.0	48.0	48.0	48.0	48.0	48.0
10^{-3} (screened/SMT-0)	72.4	5.2	0.3	0.1	0.1	0.1	0.1	0.1	0.1	0.1
10^{-3} (screened/SMT-A)	98.4	74.3	51.4	50.4	50.4	50.4	50.4	50.4	50.4	50.4
10^{-5} (not screened)	97.9	76.3	44.6	19.8	12.7	12.7	12.7	12.7	12.7	12.7
10^{-5} (screened/SMT-0)	98.2	70.9	19.1	1.1	0.2	0.1	0.1	0.1	0.1	0.1
10^{-5} (screened/SMT-A)	97.9	76.3	44.6	20.0	13.3	13.3	13.3	13.3	13.3	13.3
10^{-7} (not screened)	97.9	76.3	44.5	19.8	12.7	11.5	10.9	10.9	10.9	10.9
10^{-7} (screened/SMT-0)	97.9	76.3	44.8	20.0	5.7	0.7	0.1	0.1	0.1	0.1
10^{-7} (screened/SMT-A)	97.9	76.3	44.5	19.9	12.7	11.4	10.6	10.5	10.5	10.5
10^{-9} (not screened)	97.9	76.3	44.5	19.8	12.7	11.5	10.7	8.8	5.6	5.6
10^{-9} (screened/SMT-0)	97.9	76.3	44.5	19.8	12.7	11.4	4.6	0.4	0.1	0.1
10^{-9} (screened/SMT-A)	97.9	76.3	44.5	19.8	12.7	11.5	10.7	8.6	5.7	5.7

Table C.8: Convergence behavior of GIAO-HF/6-31G* of linear alkanes: Norm of residuum $||\text{Res}|| = ||(\text{RHS} - \text{LHS})||$.

Iteration	N_{L1}	$\text{C}_{20}\text{H}_{42}$	$\text{C}_{40}\text{H}_{82}$	$\text{C}_{80}\text{H}_{162}$	$\text{C}_{160}\text{H}_{322}$
1	19	1.2731e-01	1.2859e-01	1.2914e-01	1.2944e-01
2	15	4.1289e-02	4.1387e-02	4.1390e-02	4.1397e-02
3	10	8.6946e-03	8.8681e-03	8.9621e-03	9.0115e-03
4	6	1.2992e-03	1.3220e-03	1.3368e-03	1.3447e-03
5	2	9.0621e-04	9.2402e-04	9.4192e-04	9.4999e-04

Table C.9: Integer coefficients of EJ_5 short-ranged Schmidt-Moskowitz correlation factor [135].

Term	m	n	o
1	0	0	1
2	0	0	2
3	2	0	0
4	3	0	0
5	4	0	0

Appendix D

Atomic Units

Energy	$1 \text{ Hartree} = \mathcal{E}_a = 4.3598 \cdot 10^{-18} \text{ J}$
Mass	$m_e = 9.1095 \cdot 10^{-31} \text{ kg}$
Charge	$e = 1.6022 \cdot 10^{-19} \text{ C}$
Length	$a_0 = 5.2918 \cdot 10^{-11} \text{ m}$
Angular momentum	$\hbar = 1.0546 \cdot 10^{-34} \text{ Js}$
Electric dipole moment	$ea_0 = 8.4784 \cdot 10^{-30} \text{ Cm}$
Electric polarizability	$e^2 a_0^2 \mathcal{E}_a^{-1} = 1.6488 \cdot 10^{-41} \text{ C}^2 \text{m}^2 \text{J}^{-1}$ $= 1.4818 \cdot 10^{-25} \text{ cm}^{-3}$
First electric hyperpolarizability	$e^3 a_0^3 \mathcal{E}_a^{-2} = 3.2064 \cdot 10^{-53} \text{ C}^3 \text{m}^3 \text{J}^{-2}$ $= 8.6347 \cdot 10^{-33} \text{ cm}^5 \text{esu}$

Appendix E

Abbreviations and Symbols

AO	Atomic orbital
a.u.	Atomic units
B3LYP	Becke's three-parameter formula with Lee-Yang-Parr GGA functional
BP96(VWN)	Becke exchange functional, Perdew's correlation functional from 1986 with VWN kernel
CCA	Coupled cluster approximation
CCSD	Coupled cluster single doubles
CCSDT	Coupled cluster single doubles triples
CCSD(T)	Coupled cluster single doubles (perturbative triples)
CFMM	Continuous fast multipole method
CI	Configuration interaction
CPSCF	Coupled-perturbed self-consistent field
CSR	Compressed sparse row
D-CPSCF	Density matrix-based coupled-perturbed self-consistent field
D-QCSCF	Density matrix-based quadratically convergent self-consistent field
D-TDSCF	Density matrix-based time-dependent self-consistent field
DFT	Density functional theory

DIIS	Direct inversion in iterative subspace
DQMC	Diffusion quantum Monte Carlo
EJ ₅	Exponential short-ranged distance correlation factor
EOP	Electro-optical Pockels effect
esu	Electro-static units
eV	Electron volt
FCI	Full configuration interaction
FF	Far-field
FMM	Fast multipole method
FN	Fixed-node
FN-DQMC	Fixed-node diffusion quantum Monte Carlo
fs	Femto seconds
GGA	Generalized gradient approximation
GIAO	Gauge-including atomic orbital
HF	Hartree-Fock
HK	Hohenberg-Kohn
HOMO	Highest occupied molecular orbital
IPM	Independent particle model
K	Kelvin
KS	Kohn-Sham
LCBF	Linear combination of basis functions
LEQS	Linear equation system
LHS	Left-hand side of linear equation system
LinK	Linear exchange K
LMO	Localized molecular orbital
LSDA	Local spin-density approximation
LUMO	Lowest unoccupied molecular orbital
MAV	Maximal absolute value
MBPT	Many-body perturbation theory
MMFF94	Merck molecular force field 94
MO	Molecular orbital
MO-CPSCF	Molecular orbital-based coupled-perturbed self-consistent field

APPENDIX E. ABBREVIATIONS AND SYMBOLS

MO-TDSCF	Molecular orbital-based time-dependent self-consistent field
MP2	Møller-Plesset second order perturbation theory
NAD+	Nicotinamide adenine dinucleotide
NF	Near-field
NMR	Nuclear magnetic resonance
<i>N</i> -PDM	<i>N</i> -particle density matrix
<i>N</i> -PDM FN-DQMC	<i>N</i> -particle density matrix-based fixed-node diffusion quantum Monte Carlo
<i>N</i> -PDM VQMC	<i>N</i> -particle density matrix-based variational quantum Monte Carlo
occ	Occupied subspace
OR	Optical rectification
PIQMC	Path integral quantum Monte Carlo
PNA	Para-nitroaniline
ppm	Parts per million
ps	Pico seconds
QMC	Quantum Monte Carlo
RISSM	Row-indexed sparse storage model
RHS	Right-hand side of linear equation system
RPA	Random phase approximation
SCF	Self-consistent field
SHG	Second harmonic generation
SMT	Sparse modified thresholding
TCNQ	Tetracyano- <i>p</i> -quinodimethane
TDHF	Time-dependent Hartree-Fock
TDDFT	Time-dependent density functional theory
TDSCF	Time-dependent self-consistent field
virt	Virtual (unoccupied) subspace
VQMC	Variational quantum Monte Carlo
XC	Exchange-correlation

\AA	Ångstrom
\mathbf{A}	Vector potential
A_{oo}	Occupied-occupied projection of matrix \mathbf{A} ($A_{oo} = \mathbf{SPAPS}$, \mathbf{A} : covariant)
A_{ov}	Occupied-virtual projection of matrix \mathbf{A} ($A_{ov} = \mathbf{SPA}(\mathbf{1} - \mathbf{PS})$, \mathbf{A} : covariant)
A_{vo}	Virtual-occupied projection of matrix \mathbf{A} ($A_{vo} = (\mathbf{1} - \mathbf{SP})\mathbf{APS}$, \mathbf{A} : covariant)
A_{vv}	Virtual-virtual projection of matrix \mathbf{A} ($A_{vv} = (\mathbf{1} - \mathbf{SP})\mathbf{A}(\mathbf{1} - \mathbf{PS})$, \mathbf{A} : covariant)
$A(\mathbf{R}'_i \leftarrow \mathbf{R}_i)$	Element of acceptance matrix, acceptance probability for \mathbf{R}' emerging from \mathbf{R}
\mathbf{B}	Magnetic field vector
\mathbf{C}	Molecular orbital coefficients matrix
δ	Relative NMR chemical shifts (in ppm)
\mathbf{D}	Determinant matrix (MO)
$\tilde{\mathbf{D}}$	Inverse of transposed determinant matrix ($\tilde{\mathbf{D}} = (\mathbf{D}^\dagger)^{-1}$)
\mathbf{D}_ρ	Determinant matrix (Density)
$\tilde{\mathbf{D}}_\rho$	Inverse of determinant matrix ($\tilde{\mathbf{D}}_\rho = \mathbf{D}_\rho^{-1}$)
$\text{Dot}[\mathbf{A}, \mathbf{B}]$	Scalar product of matrices \mathbf{A} and \mathbf{B} ($\sum_{i,j} A_{ij}B_{ij}$)
\mathcal{E}	Electric field strength
E_{ecorr}	Correlation energy
E_{loc}	Local energy
E_m	Estimated energy via mixed estimator
E_T	Energy estimate
E_{xc}	Exchange-correlation energy
e^U	Correlation factor
\mathbf{F}	Fock matrix
\mathbf{F}_q	Quantum Force, drift velocity
f_{xc}	First-order derivative of XC functional
$\mathbf{G}(\mathbf{P})$	$\mathbf{G}(\mathbf{P}) = \mathbf{J}(\mathbf{P}) + \mathbf{K}(\mathbf{P})$
$G(\mathbf{R} \leftarrow \mathbf{R}', \tau)$	Green's function
g_{xc}	Second-order derivative of XC functional

\hat{H}	Hamiltonian
\mathbf{H}	Core-Hamiltonian matrix
h	Planck's constant
\hbar	Planck's constant divided by 2π
\mathbf{J}	Coulomb matrix
\mathbf{K}	Exchange matrix
\mathcal{L}	Lagrangian
L	Number of samples
$M(\mathbf{R}' \mathbf{R})$	Element of Markov matrix, transition probability for $\mathbf{R} \rightarrow \mathbf{R}'$
\mathbf{m}_A	Nuclear magnetic moment vector of atom A
$\mathcal{O}(\mathbf{M}^x)$	Scaling behavior with respect to system size
$\mathcal{O}(\mathbf{N}^x)$	Scaling behavior with respect to number of electrons
\mathcal{P}	Probability distribution
\mathbf{P}	Fock-Dirac or one-electron density in discrete representation
\mathbf{P}_{occ}	Projector onto occupied subspace ($\mathbf{P}_{\text{occ}} = \mathbf{P}\mathbf{S}$)
\mathbf{P}_{virt}	Projector onto virtual subspace ($\mathbf{P}_{\text{virt}} = \mathbf{1} - \mathbf{P}\mathbf{S}$)
\mathbf{P}^{MO}	Diagonal matrix containing occupation numbers
\mathbf{R}	Electron configuration
\mathbf{R}_A	Cartesian coordinates of nucleus A
\mathbf{r}_i	Cartesian coordinates of electron i
\mathbf{S}	Metric
\hat{T}	Kinetic energy operator
$\text{Tr}[\mathbf{A}]$	Trace of matrix \mathbf{A}
$T(\mathbf{R}'_i \leftarrow \mathbf{R}_i)$	Element of proposal matrix, proposal probability for \mathbf{R}' emerging from \mathbf{R}
thr_{SA}	Sparse algebra threshold
thr_{BF}	Basis function threshold
$\text{thr}_{\mathbf{P}}$	Compression threshold for one-particle density matrix
\mathbf{U}	Transition coefficients matrix
\hat{V}	Potential energy operator
$\mathbf{V}_{xc}^{(2)}$	Matrix representation of f_{xc}

APPENDIX E. ABBREVIATIONS AND SYMBOLS

$\mathbf{V}_{xc}^{(3)}$	Matrix representation of g_{xc}
v_{ext}	External potential
v_{eff}	Effective Kohn-Sham potential
\mathbf{x}_i	Space-spin coordinates of electron i
Z_A	Nuclear charge number of nucleus A
α	Polarizability
β	First-order hyperpolarizability
ϵ	Orbital energy
$\pi = \rho_2$	2-particle reduced density
ρ_n	n-particle reduced density
$\rho = \rho_1$	1-particle reduced density
$\hat{\rho}^\tau$	Thermal density operator
σ	Nuclear magnetic shielding tensor
σ	Standard deviation (in QMC chapter)
σ^2	Variance (in QMC chapter)
τ	Time-step size
ϕ	Molecular orbital
χ	Basis function (atomic orbital)
Ψ	Wave function
Ψ^T	Trial wave function
Ψ_{SJ}^T	Slater-Jastrow trial wave function
ω	Frequency
∇	Nabla operator

Publications

1. C. Ochsenfeld, J. Kussmann, and F. Koziol, "Ab Initio NMR Spectra for Molecular Systems with Thousand and More Atoms: A Linear Scaling Method", *Angew. Chem. Int. Ed.* **43**, 4482 (2004).
2. J. Kussmann and C. Ochsenfeld, "A Linear Scaling Implementation of the GIAO Hartree-Fock and Kohn-Sham Method for NMR Chemical Shift Calculations", in preparation.
3. Torsten Schaller, Uta P. Buechle, Frank-Gerrit Klärner, Dieter Bläser, Roland Boese, Steven P. Brown, Hans Wolfgang Spiess, Felix Koziol, Jörg Kussmann, and Christian Ochsenfeld, "The structure of molecular tweezer complexes in the solid state: NMR experiments, X-ray investigations and quantum chemical calculations", *J. Am. Chem. Soc.*, in press.
4. Wojciech Pisula, Mark D. Waston, Zeljko Tomovic, Klaus Müllen, Jörg Kussmann, Christian Ochsenfeld, Thorsten Metzroth, and Jürgen Gauss, "Helical Packing of Discotic Hexa-peri-hexabenzocoronenes: Theory and Experiment", in preparation.
5. J. Kussmann and C. Ochsenfeld, "A Linear Scaling Implementation of the TDHF and TDDFT Equations for the Calculation of Frequency-Dependent Polarizabilities and First Hyperpolarizabilities", in preparation.
6. J. Kussmann, H. Riede, and C. Ochsenfeld, "Density matrix-based variational quantum Monte Carlo providing an asymptotically linear scaling behavior for the local-energy with controllable error bounds", submitted.
7. J. Kussmann and C. Ochsenfeld, "Linear scaling fixed-node diffusion quantum Monte Carlo: Accounting for the nodal information in a density matrix-based scheme", in preparation.
8. C. Ochsenfeld, J. Kussmann, and D. S. Lambrecht, "Linear-Scaling Methods in Quantum Chemistry", in K. B. Lipkowitz and T. R. Cundari, editors, *Reviews in Computational Chemistry*, Wiley VCH, 2006 (in press).

Academic Teachers

- Prof. Dr. C. Bolm
- Prof. Dr. A. Dölle
- Prof. Dr. B. Drießen-Hölscher
- Prof. Dr. R. Dronskowski
- Prof. Dr. D. Enders
- Prof. Dr. J. Fleischhauer
- Prof. Dr. H.-J. Gais
- Prof. Dr. G. E. Herberich
- Prof. Dr. W. Hölderich
- Prof. Dr. W. Keim
- Prof. Dr. W. Leitner
- Prof. Dr. H. Lueken
- Prof. Dr. C. Ochsenfeld
- Prof. Dr. P. Paetzold
- Prof. Dr. J. Richter
- Prof. Dr. A. Salzer
- Prof. Dr. W. Stahl
- Prof. Dr. H. G. Thomas
- Prof. Dr. M. Zeidler

Curriculum Vitae

Name	Jörg Kußmann
Date of birth	12/23/1974
Place of birth	Simmern/Hsr., Germany
Parents	Rolf Kußmann, Rita Kußmann, née Billing
Citizenship	German
Marital status	single

06/2006	Ph.D. thesis with distinction (<i>summa cum laude</i>)
03/2002 - 06/2006	Ph.D. student at the Institute for Physical and Theoretical Chemistry, Prof. Dr. C. Ochsenfeld, Universität Tübingen (Germany)
12/2001 - 03/2002	Ph.D. student at the Institute for Physical Chemistry, Prof. Dr. C. Ochsenfeld, Universität Mainz (Germany)
03/2001 - 10/2001	Diploma thesis in chemistry at the Institute for Physical Chemistry, Prof. Dr. Dölle, RWTH Aachen (Germany)
10/1994 - 03/2001	Student of chemistry at the RWTH Aachen (Germany)
03/1996 - 04/1997	Zivildienst (alternative service)
06/1994	Allgemeine Hochschulreife (diploma from German secondary school qualifying for university admission or matriculation)
1984 - 1994	Herzog-Johann-Gymnasium in Simmern/Hsr.
1980 - 1984	Rottmann-Schule, Elementary school in Simmern/Hsr.